

GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: April 9, 2003, 14:27:20 ; Search time 86 Seconds  
(without alignments)  
522.305 Million cell updates/sec

Title: US-10-034-500-2

Perfect score: 1119

Sequence: 1 AEVTASCTKRKVESYNVLVDY.....QEQILVTEEVVLRGVNFAF 218

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP archaea: \*  
2: SP bacteria: \*  
3: SP fungi: \*  
4: SP human: \*  
5: SP invertebrate: \*  
6: SP mammal: \*  
7: SP mhc: \*  
8: SP organelle: \*  
9: SP phage: \*  
10: SP plant: \*  
11: SP rodent: \*  
12: SP virus: \*  
13: SP vertebrate: \*  
14: SP unclassified: \*  
15: SP virus: \*  
16: SP bacteriophage: \*  
17: SP archaea: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	109	9.7	369	16 Q8RFV5	Q8RFV5 fusobacteri
2	109	9.7	537	16 Q889A5	Q889A5 rhizobium 1
3	103.5	9.2	917	6 Q9TUB5	Q9TUB5 sus scrofa
4	99	8.8	426	16 Q8U9I9	Q8U9I9 agrobacteri
5	93.5	8.4	2857	5 Q9Y0A0	Q9Y0A0 caenorhabdi
6	92.5	8.3	615	16 Q8YK34	Q8YK34 anabaena sp
7	91.5	8.2	917	4 Q9UNF7	Q9UNF7 homo sapien
8	91	8.1	670	16 Q9L5B2	Q9L5B2 salmonella
9	90.5	8.1	347	16 Q9PHU0	Q9PHU0 xylella fas
10	90	8.0	351	3 Q03956	Q03956 saccharomyc
11	90	8.0	903	16 Q9KM24	Q9KM24 vibrio chol
12	90	8.0	903	16 Q9P9U2	Q9P9U2 xylella fas
13	90	8.0	1091	16 Q8Y4J2	Q8Y4J2 listeria mo
14	89.5	8.0	565	11 Q9CTV9	Q9CTV9 mus musculi
15	89	8.0	1084	5 Q9BP40	Q9BP40 halocynthia
16	89	8.0	2836	5 Q9U3B3	Q9U3B3 caenorhabdi

17	89	8.0	2846	5 Q9Y0A2	Q9Y0A2 caenorhabdi
18	89	8.0	2847	5 Q95X74	Q95X74 caenorhabdi
19	89	8.0	2892	5 Q9Y0A1	Q9Y0A1 caenorhabdi
20	89	8.0	2903	5 Q95X75	Q95X75 caenorhabdi
21	87.5	7.8	932	11 Q35802	Q35802 ratius norv
22	87	7.8	304	16 Q8RG85	Q8RG85 fusobacteri
23	87	7.8	2933	5 Q9GS15	Q9GS15 caenorhabdi
24	86.5	7.7	507	2 Q9AHP4	Q9AHP4 clostridium
25	86	7.6	1090	16 Q928J2	Q928J2 listeria in
26	85.5	7.6	271	16 Q8XW06	Q8XW06 clostridium
27	85.5	7.6	446	4 Q43384	Q43384 homo sapien
28	85.5	7.6	507	2 Q99Q76	Q99Q76 clostridium
29	85.5	7.6	1290	4 Q9UP26	Q9UP26 homo sapien
30	85	7.6	750	13 Q9W633	Q9W633 cyprinus ca
31	83.5	7.5	455	17 Q974G8	Q974G8 sulfobolus
32	83.5	7.5	509	4 Q75602	Q75602 homo sapien
33	83	7.4	572	10 Q9PRU2	Q9PRU2 glycine max
34	83	7.4	579	10 Q9FK07	Q9FK07 arabidopsis
35	83	7.4	1039	2 Q9F0G1	Q9F0G1 xanthomonas
36	83	7.4	1313	16 Q9PQ16	Q9PQ16 ureaplasma
37	82.5	7.4	209	16 Q9KN73	Q9KN73 vibrio chol
38	82	7.3	860	17 Q8TM25	Q8TM25 methanosa
39	82	7.3	913	11 Q9D726	Q9D726 mus musculi
40	82	7.3	1446	2 Q9F0D0	Q9F0D0 xanthomonas
41	82	7.3	1446	2 Q9EZV3	Q9EZV3 xanthomonas
42	82	7.3	2205	5 Q16857	Q16857 microcista
43	81.5	7.3	308	16 Q9ZDD3	Q9ZDD3 rickettsia
44	81.5	7.3	951	6 Q9GLY4	Q9GLY4 cycloclag
45	81.5	7.3	1548	5 Q8T9C5	Q8T9C5 dtrotophila

## ALIGNMENTS

### RESULT 1

ID	Q8RFV5	PRELIMINARY;	PRT;	369 AA.
AC	Q8RFV5;			
DT	01-JUN-2002 (TREMBLrel. 21, Created)			
DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	D-amino acid dehydrogenase large subunit (EC 1.4.99.1).			
GN	FN0576.			
OS	Fusobacterium nucleatum (subsp. nucleatum).			
OC	Bacteria; Fusobacteria; Fusobacterium.			
OX	NCBI_TaxID=76856;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 25586;			
RX	MEDLINE=2186394; PubMed=11889109;			
RA	Kapatal V., Anderson I., Ivanova N., Reznik G., Los T., Lytkidis A.,			
RA	Bhattacharyya A., Barman A., Gardner W., Grechkin G., Zhu L.,			
RA	Vasileva O., Chu L., Kogan Y., Chaga O., Goldsman E., Betnal A.,			
RA	Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,			
RA	Fonstein M., Kyriades N., Overbeek R.;			
RT	"Genome sequence and analysis of the oral bacterium Fusobacterium			
RT	nucleatum strain ATCC 25586";			
RL	J. Bacteriol. 184:2005-2018(2002).			
DR	EMBL; AE010569; ALU94772.1; -.			
KW	Oxidoreductase; Complete proteome.			
SQ	SEQUENCE 369 AA; 40851 MW; 9727259C59BDC4AF CRC64;			

Query Match 9.7%; Score 109; DB 16; Length 369;  
Best Local Similarity 25.4%; Pred. No. 0.12; 72; Indels 38; Gaps 7;  
Matches 46; Conservative 25; Mismatches

Qy	17	LVDSGSMGMGVAVREPKELAKALIKNAAPKMSYOGGLYFAFYSVILIPGSNNS	76
Db	21	VLDSSGSVTK--IGDKTMEIAKESIKVLSNP-ANAKGIVFGKGNSTSKDES	77
Qy	77	CVAECAVNTKSLDEIFG-----RLTPVG-----DGIRKHEVTINQ	112
Db	78	CGANELIYPI-GDINVEIGERALEPIOPTGWTSTIAKSIYGVEDLKALDG-----EKTIN-	131

QY 113 MPPQAAVILLDPGHHNLLGNNPVEEYKSIYQGTNPVCFHVSPADDAEGKAIIIDQVALNS 172  
 DB 132 -----ILYIITDIGITCGGNPVEIKAKOLKGEYNTIVLGIIGNVDAQNRLIKQIADAAAG 186  
 QY 173 G 173  
 DB 187 G 187  
 RESULT 2  
 Q989A5 PRELIMINARY; PRT; 537 AA.  
 ID 0989A5  
 AC 0989A5;  
 DT 01-OCT-2001 (TREMBLrel. 18, Created)  
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Hypothetical protein mlr6511.  
 GN MLR6511.  
 OS Rhizobium loci (Mesorhizobium loci).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 CC Phyllobacteriaceae; Mesorhizobium.  
 CK NCBI\_TaxID=381;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MAFF303099;  
 RX MEDLINE=21082930; PubMed=11214968;  
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
 Watanabe A., Idehara K., Ishikawa A., Kawashima K., Kimura T.,  
 Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,  
 RA Takeuchi C., Yamada M., Tabata S.;  
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
 Mesorhizobium loci";  
 RL DNA Res. 7:331-338(2000).  
 DR EMBL; AP003009; BAB52792.1; -.  
 DR InterPro; IPR002035; VWF\_A.  
 DR SMART; SM00327; VWA; 1.  
 DR PROSITE; PS50234; VMPA; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 537 AA; 56251 MW; 765768F08D59FFC6 CRC64;  
 Query Match 9.7%; Score 109; DB 16; Length 537;  
 Best Local Similarity 24.4%; Pred. NO. 0.2;  
 Matches 47; Conservative 34; Mismatches 82; Indels 30; Gaps 8;  
 QY 17 LVDYGSMMKVVAVREPKELAKAILKINAMPMSIQGLIYFPAPY----- 66  
 DB 14 ILDAAGSMAAO--IDGPKLEIARESLRTVLSVPADDEIG---FMAYGHEKSCSDI 67  
 QY 67 --VIIIFQGSWNSCVABCAVNTIKSLDIETFGRLTPVGDGIKMEHTVINGMPQAAVILLTD 124  
 DB 68 QILVPPQPSAAAILTDA-----DSLKFLGK-TPLTIAAVKQAAELKTYEDAKIVVLTID 121  
 QY 125 GHHNLGMNVEEYKSIYQGTNPVCFHVSPADDA-EGKAIIDQIVALNSGSLVDGLQLL 183  
 DB 122 GLETCGGDCALCAKEIKAGVDFTADVGFGLTADBGKI--ACLAENYGGKYL---QAS 176  
 QY 184 QNPAYCQGFVNSV 196  
 DB 177 DEKALQELALVETV 189  
 RESULT 3  
 ID 09TUB5 PRELIMINARY; PRT; 917 AA.  
 AC 09TUB5;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Epithelial chloride channel protein.  
 GN AECG.  
 OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euteheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OK NCBI\_taxid=9823;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE=ILEAL\_MUCOSA;  
 RX MEDLINE=20473747; PubMed=11015605;  
 RA Gaspar K.J., Raceite K.J., Gordon J.R., Loewen M.E., Forsyth G.W. of  
 RT "Cloning a chloride conductance mediator from the apical membrane of  
 RT porcine ileal enterocytes."  
 RL Physiol. Genomics (Online) 3:101-111 (2000) .  
 DR EMBL; AF095584; AAF00077.1; -.  
 DR InterPro; IPR000131; ATPase\_gamma.  
 DR InterPro; IPR004727; GACC\_protcl.  
 DR InterPro; IPR02035; VWF\_A.  
 DR SMART; SM00327; VMA; 1.  
 DR TIGRFAMs; TIGR00868; hCACC; 1.  
 DR PROSITE; PS00153; ATPASE GAMMA; UNKNOWN\_1.  
 DR PROSITE; PS50234; VWF\_A; 1.  
 SQ SEQUENCE 917 AA; 100735 MW; 64941944F7EAD19 CRC64;  
  
 Query Match: 9.2%; Score 103.5; DB 6; Length 917;  
 Best Local Similarity 21.4%; Pred. No. 1.2;  
 Matches 50; Conservative 39; Mismatches 78; Indels 67; Gaps 10;  
  
 QY 38 LAKAAILKINAMPMSIOGGLYTPAPSVIIIPGGSNMSCVAECANVTKEDL----- 90  
 Db 311 LDKSGSMVVGRLKRLNQAQKLFLLQ---TVEQAWGVMAFDSAAVYKSELVQINSAA 366  
 QY 91 -EIRGLTLP-----VGDIKKHKEVINOMPPQ--AAVILLTDGHN----- 128  
 Db 367 ERDALARSLPRAASGCTICGSLRAFPVYKKYPTDQSEVELLTDGDNITISACFEVK 426  
 QY 129 -----LGMNFVEBKSIVQTNPNVCFHVVSFADDAEKAIIIDIVALNSG----- 173  
 Db 427 QNGALIHVVALGPSAKKLELSQMTGLQRYA---SQGAENGLIDAFGLALSGNRAAS 483  
 QY 174 --SVLD--GIQLLNPAVCOEFPVNSVFCQEQ-----ILVTEVVVLRGVNF 216  
 Db 484 QRSIQLESQGLTLQNN-----EMNNGTVVNDSTVGKDTLFLITERKELSPIPF 532  
  
 RESULT 4  
 Q8U919 PRELIMINARY; PRT; 426 AA.  
 ID Q8U919;  
 AC Q8U919;  
 DT 01-JUN-2002 (TREMBLrel. 21, Created)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Phosphoglycerate kinase.  
 GN PGK OR ATU03739 OR AGR\_L\_2193.  
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970) .  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 CC Rhizobiaceae; Rhizobium.  
 OK NCBI\_taxid=176299;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=21608550; PubMed=11743193;  
 RX Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitzajima J.P.,  
 RA Chenu V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,  
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,  
 RA Kutyavain T., Levy R., Li M.-J., McClelland E., Palmieri A.,  
 RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,  
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krepan W., Perry M.,  
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,  
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,  
 RA Nester E.W.;  
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens  
 RT C58 " .  
 RL Science 294:2317-2323 (2001) .  
 RN [2]  
 RN SEQUENCE FROM N.A.



```

RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Quirillo B., Goldman B.S., Cao Y., Akenazi M., Halling C., Mullin L.,
RA Houmlet K., Gordon J., Vaudin M., Iartchouk O., Bpp A., Liu F.,
RA Wollam C., Allinger M., Doughly D., Scott C., Lappas C., Markelz B.,
RA Plangnon C., Crowell C., Gureon J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
DR EMBL; AF009304; AAL44549.1; ALT_INIT.
DR EMBL; AF008309; AAK89667.1; -.
KM Kinase; Complete proteome.
SQ SEQUENCE 426 AA; 44659 MW; 4CC9DA178FEBD974 CRC64;

Query Match 8.8%; Score 99; DB 16; Length 426;
Best Local Similarity 27.2%; Pred. No. 1.1;
Matches 46; Conservative 18; Mismatches 63; Indels 42; Gaps 7;

QY 65 YSVIIPQSNWSCVAECANTIKSDLEIFGRV-----TPVGDGIKHEVTINQMPQ 116
D 12 HSMVRPELATQDPRMPAFKTIIDLDNDLAGKRVLRVDLNPVADGKVTDAFRIERVAF 71
QY 117 -----AAVILTDGNNIGMNVES--VKSIVQINPNVCFHVSPADDAEGKAIIDQ 166
D 72 ILELSKCAKAYILLAHFGRPKG--EPVAMSLSQIVPTVEDVLDHAISFATDCIGAPADA 130
QY 167 IVALNSGVLVDGQLQLQN-----PAVCGE-----PNSNVF 197
D 131 VAKNDGDL-----LENTFRHKGEEKNDPFEVLELANGDIYVNDAP 174

RESULT 5
QY0A0 PRELIMINARY; PRT; 2857 AA.
ID QY0A0;
AC QY0A0;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Inositol 1,4,5-trisphosphate receptor (Hypothetical 327.2 kDa
DE protein).
GN ITR-1 OR F33D4.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilita; Rhabditoidea;
OC Rhabdilitae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=20079510; PubMed=10610772;
RA Baylis H.A., Funtlich T., Yoshikawa K., Sattelle D.J.;
RT "Inositol 1,4,5-trisphosphate Receptors are strongly expressed in the
RT Caenorhabditis elegans and are encoded by a single gene (itr-1).";
RL J. Mol. Biol. 294:467-476(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Stellyes L., Johnson D.;
RT "The sequence of C. elegans cosmid F33D4.";
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;

```

```

RT "Direct Submission.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ243182; CAB45863.1; -.
DR EMBL; AF036702; AAK68365.1; -.
DR InterPro; IPR000699; Ca-rel_channel.
DR InterPro; IPR001682; Ca/Na_pore.
DR InterPro; IPR000493; Insp3_receptor.
DR InterPro; IPR000636; Mchannel_nlg.
DR InterPro; IPR003608; MIR.
DR Pfam; PF00520; Ion_trans. 1.
DR Pfam; PF02815; MIR_2.
DR Pfam; PF01365; RYD_TPR. 2.
DR PRINTS; PR00779; INSP3RECEPTR.
DR SMART; SM00472; MIR; 3.
KM Hypothetical protein; Receptor.
SQ SEQUENCE 2857 AA; 327195 MW; 8B991F12BEC94E2F CRC64;

Query Match 8.4%; Score 93.5; DB 5; Length 2857;
Best Local Similarity 21.9%; Pred. No. 44;
Matches 52; Conservative 36; Mismatches 72; Indels 77; Gaps 12;

QY 15 NYLVDSGSMKKNVAREP-----KIELA-----KEALIKINAMPKMS----- 54
D 1921 NDKVQHSFYMKKQKDHEPEFKALTRIQTAQNRSLASDMSCSDSKPKVSVTLPLID 1980
QY 55 -----YQGLYTPAPYSVILIPQSNWSCVAECANTIKSDLEIFGRVTPVGDGIKHEET 108
D 1981 AGDTGFNGLAF-----VQGVHRHSISF--MSQLSNDLTHSIPDLAPVD-----BK 2027
QY 109 VINQMPQAAVI-----LITDGNHNGMNVBEVKSIVQINPNVCFHVSPADDAEG 160
D 2028 STDALPEVALVERILRLVQLQLCENHNSLQNFR--KQSDRTAHNLVSEITSF----- 2079

QY 161 KAIIDQVALNSGVLVDG-----LQLQNPVAVCGEFNVSVCGE 200
D 2080 ---LPTVCGSTKSLGVFGIEGHNFSLITQTALTLEFCGP--CHENQNTMAQO 2131

RESULT 6
QY0W34 PRELIMINARY; PRT; 615 AA.
ID QY0W34;
AC QY0W34;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Hypothetical protein A11782.
GN ALI1782.
OS Anabaena sp. (Strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21595285; PubMed=11759840;
RA Kameko T., Nakamura Y., Wolk C.P., Kuritz T., Sasaemoto S.,
RA Watanabe A., Iriyuchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraaki A.,
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yaouda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003587; BAB73481.1; -.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF00092; Vwf; 1.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS50234; VWF_A; 1.
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 615 AA; 66973 MW; 3586FE9D1C86018 CRC64;

Query Match 8.3%; Score 92.5; DB 16; Length 615;
Best Local Similarity 23.3%; Pred. No. 7.1;
Matches 51; Conservative 41; Mismatches 90; Indels 37; Gaps 10;

```

```

QY 1 AAVTASCTKRVESVNYLVDSGSM--MMKHVAAREPKIELAKEALIKINAMPKRSYOGG 58
DB 29 AEIPRS--PRRNILSLVIDRSGSMGAAHLHLK-----AAESV--VDQLEPKILSV 79
QY 59 LTPAPYSVIIIPQSGMNSCVAECANVTIKSDLEIFGRLLTPVDDG--IKKHETVINQMPQA 117
DB 80 VYDDADVTVPPQPTDKALKKSIRQVRA-----GGTINISGWMKGCCEVVKHOLDPOK 134
QY 118 --AVILLTDGHNHNGM--PVEEVSISYQTNPNVCFHVVSFPADDAEGKAIIDQIVALSNG 173
DB 135 INRVALLTDGHNHNGMGIQPKILATSTOKAERGITTTTLTGFGNGFEDLLIGMARAANGN 194
QY 174 -----SALVDGLQL--QNPVACOEFVNSV 196
DB 195 FYFIQSIDAAEFSEIELDSLSRVGQNLKYLELADGI 233

RESULT 7
Q9UNF7 PRELIMINARY; PRT; 917 AA.
AC Q9UNF7;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Calcium-activated chloride channel protein 2.
GN CACCC2.
OS Homo sapiens (Human).
OC Eumariata; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxId=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RX MEDLINE=99364503; PubMed=10437792;
RA Agnel M., Vermaat T., Culouscou J.M.;
RT "Identification of three novel members of the calcium-dependent
RT chloride channel (CaCC) family predominantly expressed in the
RT digestive tract and trachea.";
RL FEBS Lett. 455:295-301(1999).
DR EMBL; AF127035; AAD48398.1; -.
DR InterPro; IPR004727; CaCC_prot.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF00092; vwa; 1.
DR SMART; SM00327; VWA; 1.
DR TIGRFAMs; TIGR00868; hCaCC; 1.
DR PROSITE; PS50234; VWFA; 1.
SQ SEQUENCE 917 AA; 101153 MW; EA01090D781BEB95 CRC64;

Query Match 8.2%; Score 91.5; DB 4; Length 917;
Best Local Similarity 22.8%; Pred. No. 15;
Matches 46; Conservative 26; Mismatches 55; Indels 75; Gaps 8;

QY 17 LVDSGSMNMKHAAREPKIELAKEALIKINAMPKRSYOGGILYTPAPYSVIIIPQSGMNS 76
DB 310 VLDKSGSMG-----GKDLNRMNQAAKIFLLQ-----TVENSWSWG 345
QY 77 CV-----AECAVNTIKSDLEIFGRLLTPVG-----DGIKKHETVINQMP 115
DB 346 WHPFSTATIVKLIQIKSDERTIMAGLPY-----PLGGSISGSIKTAQVIGELMS 401
QY 116 Q---AAVILLTDGHNHNGMNPVEEVSISYQTNPNVCFHVVSFA----- 155
DB 402 QADGSEVLLITDGEDNTASSIDEVK---QSGAIVHIFALGRAADBAVIEMSKITGSHF 458
QY 156 ---DDAEGKAIIDQIVALSNGS 174
DB 459 YVSDERQNNGLIDAFGLTSGN 480

```

```

DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Orf, hypothetical protein (Hypothetical 73.9 kDa protein).
GN R0206 OR HCM1.116.
OS Salmonella typhi.
OC Plasmid R27, and plasmid pHCM1.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
CX NCBI_TaxId=601;
RN (1)
RP SEQUENCE FROM N.A.
RC PLASMID=R27;
RX MEDLINE=20280091; PubMed=10773089;
RA Sherburne C.K., Lawley T.D., Gilmour M.W., Blattner F.R., Burland V.,
RA Grobeck E., Rose D.J., Taylor D.E.;
RT "The complete DNA sequence and analysis of R27, a large IncHI plasmid
RT from Salmonella typhi that is temperature sensitive for transfer.";
RL Nucleic Acids Res. 28:2177-2186(2000).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=CT18; PLASMID=PHCM1;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davies P., Davies R.M., Dowd L., White N., Farrar P.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jorgels K.,
RA Krogh A., Larsen T.S., Leach S., Moule S., O'Goara P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrett B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
DR EMBL; AF250878; AAF70942.1; -.
DR EMBL; AL513383; CAD09713.1; -.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF00092; vwa; 1.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS50234; VWFA; 1.
KW Hypothetical protein; Plasmid; Complete proteome.
SQ SEQUENCE 670 AA; 73863 MW; 11E55C34E620E6 CRC64;

Query Match 8.1%; Score 91; DB 16; Length 670;
Best Local Similarity 25.4%; Pred. No. 11;
Matches 44; Conservative 26; Mismatches 77; Indels 26; Gaps 8;

QY 19 DYSGSMNMKHAAREPKIELAKEALIKINAMPKRSYOGGILYTPA--PYSVIIPQSGMNS 76
DB 506 DISGSMSEHRYIHAIKTDLALTAIE--AIKRNHVNVIYPVDKDFEVI-----K 555
QY 77 CVACAVNTI--KSDLEIFGRLLTPVDDGIIKMH--ETVINQMPQAAVILLTDGHNHNGMNPV 134
DB 556 TFDENAEKTLSTFSGCGKGNNTPTGSALNMALELLBEOFRKIVFLITDGPYTSAYTI 615
QY 135 EEVKSISYQTNPNVCFHVVSFPADDAEGKAI-IDQIVALSNGS-VLVDGLQLLN 185
DB 616 NDVFVASEN-----GIEIAGVGIKTVLMGFNEGFEVNVNDISLIPN 658

```

```

RESULT 8
Q9L5E2 PRELIMINARY; PRT; 670 AA.
ID Q9L5E2
AC Q9L5E2;

```

```

Q9PHD0 PRELIMINARY; PRT; 347 AA.
ID Q9PHD0
AC Q9PHD0;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Conjugal transfer protein.
GN XFA0015.
OS Xylella fastidiosa.
OC Plasmid pXF51.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.

```

OX NCBI\_TaxID=2371;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=9A5C;  
 RX MEDLINE=20365717; PubMed=10910347;  
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,  
 Alvarenga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,  
 Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Bittes M.R.S.,  
 Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carver H.,  
 Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.W.,  
 Coutinho L.L., Cristofani M., Dias-Neto B., Docena C., El-Dorri H.,  
 Falcinani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,  
 Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,  
 Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,  
 Ho P.L., Hohlseil J.D., Jungman M.L., Kemper E.L., Kikijima J.P.,  
 Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,  
 Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,  
 Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Martino C.L.,  
 Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,  
 Marck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,  
 Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.R.S.,  
 Nhami A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
 de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,  
 Peixoto B.R., Pereira G.A.G., Pereira H.A.Jr., Pesquero J.B.,  
 Queiroz R.B., Roberto P.G., Rodrigues V., de Rosa A.U.M.,  
 da Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.B.,  
 da Silva A.C.R., da Silva A.M., da Silva P.R., Silva W.A. Jr.,  
 da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,  
 de Souza A.P., Terenzi M.F., Truffi D., Testi S.M., Teuhako M.H.,  
 Vallada H., Van Slyke M.A., Verjovski-Almeida S., Vettore A.L.,  
 Zago M.A., Zatz M., Zeldman J., Zetubal J.C.;  
 RA "The genome sequence of the plant pathogen *Xylella fastidiosa*."  
 RT Nature 406:151-159(2000).  
 RL EMBL: AE003851; AAF8584.1;  
 DR InterPro: IPR001482; GSPIT\_E.  
 DR Pfam: PF00437; GSPIT\_E; 1.  
 DR ProDom: PD000739; GSPIT\_E; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 347 AA; 38430 MW; D1012D3336C5D9F0 CRC64;  
 Query Match 8.1%; Score 90.5; DB 16; Length 347;  
 Best Local Similarity 22.1%; Pred. No. 5.1;  
 Matches 47; Conservative 35; Mismatches 80; Indels 51; Gaps 9;  
 QY 18 VDYGSGMMKQVAVREPKIEIAKEAILKINAMPKMSYGGIYTFAPYSVIIPOGSMNC 77  
 DB 144 VDSGTLHRYVEL-----LALKSKRIK--EFLKLAQHG-----TLVIYKTSQSKTT 192  
 QY 78 VAECVANTIKSDLEIFGRITPVGDGIKNHETVINQMPQAAVILLTDGHNNLGNPVEEV 137  
 DB 193 IGSITNCIPTD---EELVTVED---VHEMFLNHPKKNLHFRSDDEGSLINPKQAI 245  
 QY 138 KSIYOTNNVCGHVSFDDA-----EGKAI--IQIYVANSGLV 177  
 DB 246 ASCLRMKEDRILLTEMRGDEMEFVKAVGSGHPGISHTAGALRAFDIYALIDS--A 303  
 QY 178 DGLQLQNPVAVCOEFNVSCFOEILTVTEVVV 210  
 DB 304 TGAHL-----DAAIYKKRVETVDIVL 325  
 RESULT 10  
 Q9KIM24 PRELIMINARY; PRT; 318 AA.  
 AC O9KIM24;  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
 DE Hypothetical protein VCA0172.  
 GN VCA0172.  
 OS Vibrrio cholerae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrrio.  
 NCBI\_TaxID=666;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=EL TOR N16961 / SEROTYPE O1;  
 RX MEDLINE=20406833; PubMed=10952301;  
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,  
 Dodson R.J., Hatt D.H., Hickey E.K., Peterson J.D., Umayam L.A.,  
 Gill S.R., Nelson K.E., Read T.D., Tetselin H., Richardson D.,  
 Ermlaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,  
 McDonald L., Uterback T., Fleischmann R.D., Nieman W.C., White O.,  
 Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
 Fraser C.M.;  
 RA "DNA sequence of both chromosomes of the cholera pathogen *Vibrio*  
 RT cholerae."  
 RT Nature 406:477-483(2000).  
 RL EMBL: AE004357; AAF6085.1;  
 DR InterPro: IPR002035; VWF\_A.  
 DR Pfam: PF00092; vwa; 1.  
 DR SMART: SM00327; VWA; 1.  
 DR PROSITE: PS50234; VWFA; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 318 AA; 35753 MW; 06565A71FA86C6252 CRC64;  
 Query Match 8.0%; Score 90; DB 16; Length 318;  
 Best Local Similarity 26.7%; Pred. No. 5;  
 Matches 36; Conservative 22; Mismatches 53; Indels 24; Gaps 6;  
 QY 58 GLYTFAPYSVI--IQGSMNSGVAECVANTIKSDLEIFGRITPVGDGIKNHETVINQMP 115  
 DB 131 GLIFPADVAIYQTPITLTDQTVANQNGTV---LKLIGTQTAIGEGIGLAKTPTFIDSDAP 187  
 QY 116 QAAVILLTDGHNNIC-KNPVEVKSIVQTNPNV-----CFHVSFADDA 158  
 DB 188 QRWILLSDGSGNTAGVDPLPAAINAIKQYNTTIVYGVGAGEMVYKDLFSKRVTAQDL 247  
 QY 159 EGKAIIDQIVANSG 173  
 DB 248 DEKTL--QTIAVTG 260  
 RESULT 11  
 Q03956 PRELIMINARY; PRT; 351 AA.  
 AC Q03956;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE Hypothetical 40.0 kDa protein.  
 GN RAV2 OR YDR202C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AB972;  
 RA Oliver K., Harris D.;  
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AB972;  
 RA Bartell B., Rajandream M.A., Walsh S.V.;  
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: Z48784; CA88714.1; --.  
 DR SGD: S0002610; RAV2.  
 KW Hypothetical protein.  
 SQ SEQUENCE 351 AA; 40017 MW; 0194C41FBF3819E9 CRC64;  
 Query Match 8.0%; Score 90; DB 3; Length 351;  
 Best Local Similarity 20.3%; Pred. No. 5.7;  
 Matches 52; Conservative 35; Mismatches 69; Indels 100; Gaps 12;  
 QY 16 YLVYSGMMKQVAVREPKIEIAKEAILKINAMPKMSYGGIYTFAPYSVIIPOGSMN 75

```

Db 90 YIVDP-----HIVRPPQFGKQVMPRMVGTGLNPLLIQ-----FSKIMTH----- 130
Oy 76 SCVACAVNTKSDLEIFGRILTPVD-----GIKM-----HEVINOQMPQAAVILLT 123
Db 131 -----LKNILIEINLOQVATVSEFVSKFQVAMELNHSLILLQNPDRD---LVF 177
Oy 124 DGHNNLGNPVEEVSIVOTNPNC---FHVVSF----- 154
Db 178 PEDNNFAM-----KEMQDCYSCVCESTRHILIGLEITLCRNELCTELRNILKVTKKPNC 231
Oy 155 ADDEAGKAIIDQI-----VALNCGSVLVYDGLQILLQNPVCOEFVNSVFCQEQ 201
Db 232 IDSKGRSFCQDIRNQVTERNKTKLSIKISENGVQVQDSTLLNH-----IISFQSEA 284
Oy 202 ILVTE-EVVLVLRGVNF 216
Db 285 ITPFAQELLRGVTF 300

RESULT 12
Oy 909P9U2 PRELIMINARY: PRT: 903 AA.
AC 09P9U2;
Dt 01-OCT-2000 (TREMBLrel. 15, Created)
Dt 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
Dt 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
De Hypothetical protein Xf2779.
GN Xf2779.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
CX Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9ASC;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Bala G.S., Battista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Britones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carter H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.P., Cristofani M., Dias-Neto E., Docena C., El-Dorty H.,
RA Facchini A.L., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,
RA Garnier J.M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hohnes J.D., Junqueira M.L., Kemper E.L., Kitajima U.P.,
RA Krieger J.B., Kuramae E.B., Laigret F., Lampais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Martino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Nham D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nham A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Queiroz R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Sanceli R.V., Sasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
RA Vallada H., Van Sluys M.A., Verjovsky-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Zeldin J., Zetser J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa";
RL Nature 406:151-159 (2000).
DR EMBL; AE004083; AAF85564.1; -.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR002509; Polysac_deacet.
DR InterPro; IPR001440; TPR.
DR Pfam; PF01522; Polysac_deacet; 1.
DR PROSITE; PS00018; EF_HAND; UNKNOWN.1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 903 AA; 101108 MW; E0A31EC7A6A1A8CC CRC64;

```

```

Query Match 8.0%; Score 90; DB 16; Length 903;
Best Local Similarity 20.9%; Pred. No. 20;
Matches 54; Conservative 45; Mismatches 101; Indels 58; Gaps 10;

Oy 1 AETVASTKRVESNYLVDSGSM-----MKHVAREPKIELAEALIKINAAMPKSYQ 56
Db 112 ALLASGAPSRDPTLVLLDYIESSPDLYDADRLEFRELNDL--DARINHGALPAIKLR 169
Oy 57 GGLV-----TFAPY-----SVIIPGNSVCVACAVN-TIKSDLEI 92
Db 170 CRHEDDALNQIERNNQETRLSLIPDRNRGIFIKERKDDVVAHNRSTYTKMLIHD 229
Oy 93 FGRILT--FVGDGIKHEVINOQMPQAAVILLTDGHNNLGNPVEEVSIVOTNPNCFH 150
Db 230 YGILEAVYIPMESESHETFGNELPPKTLVLTFFDDGPHRYVNEIKELIIOHYAV-PAVFFE 288
Oy 151 VVSFAD--DAEGK-----AIIQDVAIANSGLVYDGLQILLQNPVCOEFVNSVF 197
Db 289 VGRNIGRFRDRAKPOLGPLSKITRELIQGYAVVANSHTDLSKLSGNALKEVIN-- 346
Oy 198 COEQILVTEEVVLVLRGVN 215
Db 347 -----DILRAVD 354

RESULT 13
Oy 08Y4J2 PRELIMINARY: PRT: 1091 AA.
AC 08Y4J2;
Dt 01-MAR-2002 (TREMBLrel. 20, Created)
Dt 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
Dt 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
De Hypothetical protein lmo2446.
GN lmo2446.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
CX Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ESD-E / SEROVAR 1/2A;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangoul L., Buchrieser C., Rusniok C., Amend A.,
RA Baguerio F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Checrouni F., Couve E., de Darvar P., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussargat O.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Haut J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kutepkat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Weiland U., Coesart P.;
RT "Comparative genomics of Listeria species";
RL Science 294:849-852 (2001).
DR EMBL; AL591983; CAD00524.1; -.
DR ListList; lmo02446; -.
DR InterPro; IPR000322; Glyco_hydro_31.
DR Pfam; PF01055; Glyco_hydro_31; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 1091 AA; 121698 MW; 55FC3969FA72DCDE CRC64;

Query Match 8.0%; Score 90; DB 16; Length 1091;
Best Local Similarity 22.0%; Pred. No. 25;
Matches 56; Conservative 37; Mismatches 74; Indels 88; Gaps 15;

Oy 21 GSGMMKHVAVEPKIELAKEALIKINAAMPKSYQ--GGLYTFAPY-----SVIIPQ 71
Db 743 NGKSMKRWQAMMYPE-----DYNARDLDEQIMFGDDLVAIPVSGQTEKEVLYPE 793
Oy 72 GSNWSCVACAVNTKSDLEIF--GRILTPVGDGIKHEVINOQMP--PQAAVIL---LTDG 125

```

Db 794 GEM-----VDIMNGVHPGGETISTYADV-DTLFVPAKAGAIIPMMNTDG 837

Qy 126 HNNIGMNPVEEVKSIYQTNPVVCFHV-----VSFADDAEG-----XAI 163

Db 838 Y-QGQGVNNDLKS-----DMLTFVYPSGDSSEYFYPDDVNGEMRDISVSEDFANEKYS 892

Qy 164 ID-----QIVALNSGYLVNGLQLLQ-----NPVCOEFVNSV---FCQBOI 202

Db 893 VDLPMADETTMQVSTPTSTLIDGADVAKADTLDAFNEATGYGYDTQVLTLYVKAAL 952

Qy 203 LVEEVVVLRGVNPFA 217

Db 953 KDQKQATLVNGVNH 967

RESULT 14

Q9CTV9 PRELIMINARY; PRT; 565 AA.

AC Q9CTV9

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE 5830475106Rik protein (Fragment).

GN 5830475106Rik.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI\_TaxId=10090;

[1]

RN SEQUENCE FROM N.A.

RC STEAIN=C57BL/6J; TISSUE=PIUTITARY GLAND;

RC MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Arikawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Aizawa K., Iwawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I., Saito T., Okazaki Y., Gojopori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Peeble G., Quackenbush J., Schirral L.M., Steubli F., Suzuki R., Tomita M., Wagner L., Washio T., Sekai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., R. Guentrich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombereis P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmink L., Wyshaw-Boris A., Yoshida K., Haesegawa Y., Kawaji H., Kohlsuki S., RA Hayaishizaki Y.;

RT "Functional annotation of a full-length mouse cDNA collection."

RL Nature 409:685-690(2001).

DR EMBL; AK019914, BAB31913.1, -

DR MGD; MGI:1915026, 5830475106Rik.

DR InterPro; IPR002035; VWF\_A.

DR SMART; SM00327; VWA; 1.

DR PROSITE; PSS0234; VWF\_A; 1.

FT NON TER 565

SO SEQUENCE 565 AA; 62297 MW; E502E7C1AC2FC08C CRC64;

Query Match 8.0%; Score 89.5; DB 11; Length 565;

Best Local Similarity 27.9%; Pred. No. 12;

Matches 55; Conservative 31; Mismatches 88; Indels 23; Gaps 11;

Qy 2 EVTASCTRVESYNYLVYSGM--MKKVAVREPKIELAKAIIKINAMPKNGY----- 55

Db 272 EVKAS--TACEFVFLPMRSGSDMSPSTENNSQRIEAKETLLLLKSLPMGCFNYI 329

Qy 56 -QGGLY-TEAPYSVILIPQSNNSCVAECAVNTIKSDLIFGRLTPVGDIKHETVINGM 113

Db 330 GFGSSYEKFPFESVYKYDTDMEDAVK--VKALKANLGGTELLTPLCNLYK--ASSIFGH 385

Qy 114 PQAAVILLTDGHNHNGMNPVEEVKSIYQTNPVVCFHVNSFA--DDAEGKAIIDQIVALNS 172

Db 386 PLQ--LFLVFTDGEVSDTRFVSIREVK--LNSKXRCF-----SFGIGGASTSLIKNIARVSG 438

Qy 173 G-SVLVDGLQLLONPAV 188

Db 439 GTAVFITKDMQTKAL 455

RESULT 15

Q9BP40 PRELIMINARY; PRT; 1084 AA.

AC Q9BP40

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE Complement factor B.

OS Halocynthia roretzi (Sea squirt).

OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Stolidobranchia; Pyuridae; Halocynthia.

NCBI\_TaxId=7729;

[1]

RN SEQUENCE FROM N.A.

RC TISSUE=HEPATOPANCREAS;

RA Ji X., Nankawa-Yamada C., Nakanishi M., Sasaki M., Nonaka M.;

RT "Unique domain structure of ascidian complement factor B: Trace of exon shuffling".

RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPsin FAMILY.

DR EMBL; AF224491; AAK00631.1; -

DR HSSP; P00763; IDPO.

DR InterPro; IPR001314; Chymotrypsin.

DR InterPro; IPR002172; LDL\_recept\_A.

DR InterPro; IPR001254; Ser\_protease\_Try.

DR InterPro; IPR000436; Sush1\_SCR\_CCF.

DR InterPro; IPR002035; VWF\_A.

DR Pfam; PF00084; trypsin; 5.

DR Pfam; PF00089; trypsin; 1.

DR Pfam; PF00092; vwa; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR PRINTS; PR00261; LDLRECEPTOR.

DR SMART; SM00032; CCP; 5.

DR SMART; SM00192; LDLA; 1.

DR SMART; SM00020; TRYP\_SPC; 1.

DR SMART; SM00327; VWA; 1.

DR PROSITE; PS01209; LDLRA\_1; 1.

DR PROSITE; PS50068; LDLRA\_2; 1.

DR PROSITE; PS50240; TRYPsin\_DOM; 1.

DR PROSITE; PS00135; TRYPsin\_SER; 1.

DR PROSITE; PSS0234; VWF\_A; 1.

DR GlycoProtein; Hydrolase; Serine protease.

SO SEQUENCE 1084 AA; 121213 MW; 4FEC8109E7B00CC CRC64;

Query Match 8.0%; Score 89; DB 5; Length 1084;

Best Local Similarity 20.9%; Pred. No. 31;

Matches 46; Conservative 36; Mismatches 70; Indels 68; Gaps 9;

Qy 2 EVTASCTRVESYNYLVYSGM--MKKVAVREPKIELAKAIIKINAMPKNGYGGLYT 61

Db 536 DVSKSVTKYDDFS-----SG-----IEFAKRLI-----DRKNKGVLK 570

Qy 62 PAPYSVILIPQSNNSCVAECAVNTIKSDLIFGRLTPVGDDGK----- 104

Db 571 YS-----IIAYASNKTQLEITDRSTNVKVIKRLDNDSDQVKAIVSELIEFTNSGATA 626

Qy 105 -----MKETVINGMPQAAVILLTDGHNHNGMNPVE--EVKSIYQTNPVNC 148

Db 627 KALKSLRMLFMFEMDINDQTNKCHVFLFTDGNHNGMNPVEVRKMKQKIF--GSNIE 684

Qy 149 FHVVSFADDAEGKAIIDQIVALNS--GSVLVDGLQLLON 185

Db 685 FYSISAQEDPSPFA--FELLIGLASEPENVIYIEDIHLSS 723

Thu Apr 10 07:47:19 2003

us-10-034-500-2.rspt

Page 8

Search completed: April 9, 2003, 14:32:15  
Job time : 92 secs

---

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 9, 2003, 14:18:41; Search time 25 Seconds  
(without alignments)  
361.674 Million cell updates/sec

Title: US-10-034-500-2

Perfect score: 1119  
Sequence: 1 AEVTASCTKRVESYNVLVDY.....QEOILVTEEVVLRGVNFAP 2:8

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query	Match Length	ID	Description
1	126	11.3	227	1 YWMC_BACSU	P70960 bacillus su
2	106.5	9.5	224	1 YWMD_BACSU	P70961 bacillus su
3	92	8.2	885	1 ITH3_HUMAN	Q06033 homo sapien
4	91	8.1	233	1 TNFA_MARMO	O35734 marocla mon
5	90	8.0	459	1 DHE4_EMENI	P18819 emericella
6	84	7.5	271	1 MURI_MYCTU	Q10626 mycobacteri
7	81.5	7.3	398	1 RCG_RHIL0	Q086f1 rhizobium l
8	81.5	7.3	1164	1 YAVI_XANCV	P14727 xanthomonas
9	81	7.2	886	1 ITH3_MOUSE	Q61704 mus musculu
10	80.5	7.2	358	1 CADH_MEUSA	P31656 medicago sa
11	80	7.1	391	1 GPDI_YEAST	O00055 saccharomyc
12	80	7.1	753	1 YV79_SCHPO	O09763 echinosacch
13	79.5	7.1	1165	1 RRC2_SCHRO	Q10233 echinosacch
14	79.5	7.1	177	1 IDI_RHOSH	Q095d3 rhodobacter
15	79	7.1	486	1 THRC_HELPY	O24924 hellicobacte
16	79	7.1	601	1 ISPG_CHLMU	Q04931 chlamydia m
17	79	7.1	875	1 GYRA_ECOLI	P09097 escherichia
18	79	7.1	878	1 GYRA_SALTY	P37411 salmonella
19	79	7.1	2183	1 RRP1_MEASA	P35975 measles vir
20	78.5	7.0	317	1 ROH3_RAT	P05306 rattus norv
21	78.5	7.0	335	1 YE81_MYCAV	O07395 mycobacteri
22	78.5	7.0	610	1 PTMA_BACSU	P42956 bacillus su
23	78.5	7.0	840	1 GYRA_UREPA	O09r63 ureaplasma
24	78	7.0	417	1 PGK_CAEEL	P06416 caenorhabdi
25	78	7.0	887	1 ITH3_RAT	O64416 rattus norv
26	78	7.0	1460	1 PMPC_CHLMU	Q09jy1 chlamydia m
27	77.5	6.9	332	1 YC68_MYCTU	Q11051 mycobacteri
28	77.5	6.9	317	1 ROH1_RAT	P50169 rattus norv
29	77.5	6.9	486	1 THRC_HELPY	O24924 hellicobacte
30	77	6.9	226	1 TRIS_METFR	O74025 methanobact
31	77	6.9	286	1 PYRF_YARLI	Q12724 yarrowia li
32	77	6.9	321	1 GIK_ECOLI	P46880 escherichia
33	77	6.9	735	1 CIGB_DICDI	Q04481 dictyosteli

34	77	6.9	876	1 GYRA_KLEPN	P14829 klebsiella
35	76.5	6.8	903	1 ECLC_BOVIN	P54281 bos taurus
36	76.5	6.8	1644	1 RPA1_DROME	P18185 drosophila
37	76	6.8	322	1 GCA_RAT	P20760 rattus norv
38	76	6.8	426	1 NODC_RHIME	P04341 rhizobium m
39	76	6.8	633	1 YR45_CAEEL	O09563 caenorhabdi
40	76	6.8	663	1 ALOX_CANBO	O00922 candida boi
41	76	6.8	719	1 PBPA_STRPN	Q04707 streptococc
42	76	6.8	756	1 MAO2_HABIN	P43837 haemophilus
43	76	6.8	2183	1 RRP1_MEASE	P12576 measles vir
44	75.5	6.7	338	1 FLIG_BACSU	P23448 bacillus su
45	75.5	6.7	501	1 URK1_YEAST	P27515 saccharomyc

## ALIGNMENTS

RESULT 1	YWMC_BACSU	STANDARD;	PRT;	227 AA.
ID	YWMC_BACSU			
AC	P70960;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Hypothetical protein ywmc precursor.			
GN	YWMC.			
OS	Bacillus subtilis.			
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.			
OX	NCBI_TaxID=1423;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=168;			
RX	MEDLINE=98045417; PubMed=9384377;			
RA	Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,			
RA	Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,			
RA	Bourllet S., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,			
RA	Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,			
RA	Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,			
RA	Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,			
RA	Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,			
RA	Ghim S.Y., Glaeser P., Goffeau A., Golightly E.J., Grandi G.,			
RA	Guiseppi G., Guy B.J., Haga K., Haech U., Harwood C.R., Henaut A.,			
RA	Hilbert H., Holstappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,			
RA	Joris B., Karamata D., Kasehara Y., Kleier-Blanchard M., Klein C.,			
RA	Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,			
RA	Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,			
RA	Lee S.M., Levine A., Liu H., Maeda S., Manuel C., Medigue C.,			
RA	Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Nodack M.,			
RA	Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudaga B., Park S.H.,			
RA	Paro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,			
RA	Pressean E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,			
RA	Rieger M., Rivoita C., Roche E., Roche B., Rose M., Sadle Y.,			
RA	Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,			
RA	Sekiguchi J., Sekowsky A., Seror S.J., Seror P., Shin B.S., Solito B.,			
RA	Sotokun A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,			
RA	Takenuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,			
RA	Tosato V., Uchiyama S., Vandenbol M., Vamier F., Vaasaroti A.,			
RA	Viari A., Wambut R., Wedler E., Wedler H., Weitzenecker T.,			
RA	Winters P., Wiput A., Yamamoto H., Yamane K., Yamamoto K., Yata K.,			
RA	Yoshida K., Yoshikawa H.F., Zumschein E., Yoshikawa H., Danchin A.,			
RT	"The complete genome sequence of the Gram-positive bacterium Bacillus subtilis."			

Nature 390:249-256(1997).  
 CC -1- SIMILARITY: STRONG, TO B.SUBTILIS YWMD.  
 CC -1- SIMILARITY: CONTAINS 1 WMFA DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements/>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 CC EMBL; 281356; CAB03680.1; -  
 CC DR EMBL; 299122; CAB15690.1; -  
 CC DR Subtilist; BG12474; YWMD.  
 CC DR InterPro; IPR002035; VWF\_A.  
 CC Pfam; PF00092; VWA; 1.  
 CC DR SMART; SM00327; VWA; 1.  
 CC DR PROSITE; PS50234; WMFA; 1.  
 CC KW Hypothetical protein; Signal; Complete proteome.  
 CC FT SIGNAL 1 23  
 CC FT CHAIN 1 23  
 CC FT DOMAIN 36 227  
 CC FT SEQUENCE 227 AA; 24542 MW; ACB1D2A106310A5E CRC64;  
 SO  
 Query Match 11.3%; Score 126; DB 1; Length 227;  
 Best Local Similarity 24.6%; Pred. No. 0.00032;  
 Matches 57; Conservative 37; Mismatches 84; Indels 54; Gaps 11;  
 OY 1 AEVTASCTRVSYNYLVYDSSGMMKHAVAVERPKIELAKEILKINAMP-----KMSY 55  
 DB 24 AEKTEAKAPANVAIVLIDASSGMARIDGV--SKFNASKKEISKFPASSLPSTGYKMSV 81  
 OY 56 QG-----GLTFAPYRVIIIPQGSWNSCVAECAANTIKSLDEIFGRIT 97  
 DB 82 FSESGNNKSKVQSCGSAIRNYGFGFEN--EQSFLNS-----LNTIGPT-----GMT 127  
 OY 98 PVDGDIKMEETVJINQMP--QAAVILLTDGHNNLIGNPVEEVKSIYQTPNVCFHVVSFA 155  
 DB 128 PLAKALNEKSSFPDQDAGEKVVYLLTDGEETGCGNPIKTKALEK--DNITVNVIGF- 184  
 OY 156 DPAEG-KAIIIDQIVAINSGSVLVGLQILQNPVACQEPVNSVFCQEQILVTE 206  
 DB 185 DYKEGKQQLNAIKVGGGEYF-----PAYTQKDYEKIFTOQSLMLSK 227  
 RESULT 2  
 YWMD BACSU STANDARD; PRT; 224 AA.  
 AC P70961;  
 DT 16-OCT-2001 (Rel. 40; Created)  
 DT 16-OCT-2001 (Rel. 40; Last sequence update)  
 DT 15-JUN-2002 (Rel. 41; Last annotation update)  
 DE Hypothetical protein ywmd precursor.  
 GN YWMD.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RA MEDLINE=98015417; PubMed=9353933;  
 RA Presecan E., Moszer I., Boursier L., Cruz Ramos H.C., De La Fuente V.,  
 RA Hullo M.-F., Lelong C., Schleich S., Sekowska A., Song B.H.,  
 RA Villani G., Kunst F., Danchin A., Glaeser P.,  
 RT "The Bacillus subtilis genome from gerBC (311 degrees) to llyC (334  
 RT degrees).";  
 RL Microbiology 143:3313-3328(1997).  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=98044033; PubMed=9384377;  
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,

Azevedo V., Bertero M.G., Beesieres P., Bolotin A., Borchert S.,  
 RA Boursier L., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
 RA Bouillet S., Brunech C.V., Caldwell B., Capuano V., Cater N.M.,  
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
 RA Denizot F., Devine K.M., Dusterhoft A., Ehlich S.D., Emerson P.T.,  
 RA Entian K.D., Errington J., Fabre C., Ferrari E., Foulger D.,  
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizzi A., Galleron N.,  
 RA Gilm S.Y., Glaeser P., Goffeau A., Golightly E.J., Grandi G.,  
 RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,  
 RA Hilbert H., Holnappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
 RA Joris B., Karamata D., Kasahara Y., Klarrer-Blanchard M., Klein C.,  
 RA Kobayashi Y., Koester P., Koningstein G., Krogh S., Kumano M.,  
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
 RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,  
 RA Medina N., Mellado R.P., Mizuno M., Moesli D., Nakai S., Noback M.,  
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
 RA Parro V., Pohl T.M., Portetelle D., Portwollik S., Prescott A.M.,  
 RA Presecan E., Puig P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
 RA Riesger M., Rivolta C., Roche B., Rose M., Sadale Y.,  
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
 RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,  
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
 RA Takeuchi M., Tanakoshi A., Tanaka T., Terpetra P., Tognoni A.,  
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassart A.,  
 RA Viari A., Wambut R., Wedler E., Wedler H., Weitzenecker T.,  
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
 RA Yoshida K., Yoshikawa H.F., Zumberg E., Yoshikawa H., Danchin A.,  
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus  
 RT subtilis.";  
 RL Nature 390:249-256(1997).  
 CC -1- SIMILARITY: STRONG, TO B.SUBTILIS YWMD.  
 CC -1- SIMILARITY: CONTAINS 1 WMFA DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements/>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 CC EMBL; 281356; CAB03680.1; -  
 CC DR EMBL; 299122; CAB15690.1; -  
 CC DR Subtilist; BG12475; YWMD.  
 CC DR InterPro; IPR002035; VWF\_A.  
 CC DR SMART; SM00327; VWA; 1.  
 CC DR PROSITE; PS50234; WMFA; 1.  
 CC KW Hypothetical protein; Signal; Complete proteome.  
 CC FT SIGNAL 1 23  
 CC FT CHAIN 24 224  
 CC FT DOMAIN 31 224  
 CC FT SEQUENCE 224 AA; 23965 MW; 4614AC7F9025E397 CRC64;  
 SO  
 Query Match 9.5%; Score 106.5; DB 1; Length 224;  
 Best Local Similarity 21.9%; Pred. No. 0.017;  
 Matches 42; Conservative 29; Mismatches 66; Indels 55; Gaps 9;  
 OY 17 LVYDSSGMMKHAVAVERPKIELAKEILKINAMP-----MSYGG----- 58  
 DB 35 LFDGSSGWQK--TGGERKIDIAKKSVAFAELLPQDNTMLRVGCHGNNKLSKALASC 92  
 OY 59 -----LYTFAPY-----SVIIPQGSWNSCVAECAANTIKSLDEIFGRITPVGDGIK 104  
 DB 93 STTETITGHPYEGSLFNSSELPRTG-WTP-IAKALADT-RKFPFAFD-----ADG-- 142  
 OY 105 MEETVJINQMPQAAVILLTDGHNNLIGNPVEEVKSIYQTPNVCFHVVSFAADAGKAIT 164  
 DB 143 -----KNVYLLTDGEETGCGDPAEIEKLRASNVDTIYVITIGFNPDVGNBEM 191  
 OY 165 DQIVAINSGSVL 176  
 DB 192 KQAAVAGGEYI 203



RESULT 3  
ID ITIH3 HUMAN STANDARD; PRT; 885 AA.  
AC 006033; 099085;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-OCT-2001 (Rel. 40, Last annotation update)  
DE Inter-alpha-trypsin inhibitor heavy chain H3 precursor (ITI heavy chain H3) (serum-derived hyaluronan-associated protein) (SHAP).  
GN ITIH3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OC NCBI\_Taxid=9606;  
[1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=93215656; PubMed=7681778;  
RA Bourguignon J., Diarra-Mehrpour M., Thiberville L., Bost F., Seeboue R., Martin J.P.;  
RT "Human pre-alpha-trypsin inhibitor-precursor heavy chain. cDNA and deduced amino-acid sequence.";  
RL Eur. J. Biochem. 212:771-776(1993).  
[2]  
RP SEQUENCE OF 341-885 FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=89137072; PubMed=2465147;  
RA Diarra-Mehrpour M., Bourguignon J., Seeboue R., Mattei M.-G., Passage E., Saller J.P., Martin J.P.;  
RT "Human plasma inter-alpha-trypsin inhibitor is encoded by four genes on three chromosomes.";  
RL Eur. J. Biochem. 179:147-154(1989).  
[3]  
RP SEQUENCE OF 30-49; 463-477 AND 497-515.  
RX MEDLINE=89380192; PubMed=2476436;  
RA Englid J.J., Thøgersen I.B., Pizzo S.V., Salvesen G.;  
RT "Analysis of inter-alpha-trypsin inhibitor and a novel trypsin inhibitor, pre-alpha-trypsin inhibitor, from human plasma.";  
RL J. Biol. Chem. 264:15975-15981(1989).  
[4]  
RP SEQUENCE OF 631-647, AND CROSS-LINKAGE SITE TO BIKUNIN.  
RX MEDLINE=91093267; PubMed=1898736;  
RA Englid J.J., Salvesen G., Hetta S.A., Thøgersen I.B., Rucheford S., Pizzo S.V.;  
RT "Chondroitin 4-sulfate covalently cross-links the chains of the human blood protein pre-alpha-inhibitor.";  
RL J. Biol. Chem. 266:747-751(1991).  
-1- FUNCTION: MAY ACT AS A CARRIER OF HYALURONAN IN SERUM OR AS A BINDING PROTEIN BETWEEN HYALURONAN AND OTHER MATRIX PROTEIN, INCLUDING THOSE ON CELL SURFACES IN TISSUES TO REGULATE THE LOCALIZATION, SYNTHESIS AND DEGRADATION OF HYALURONAN WHICH ARE ESSENTIAL TO CELLS UNDERGOING BIOLOGICAL PROCESSES.  
-1- SUBUNIT: I-ALPHA-1 PLASMA PROTEASE INHIBITORS ARE ASSEMBLED FROM ONE OR TWO HEAVY CHAINS (H1, H2 OR H3) AND ONE LIGHT CHAIN, BIKUNIN. INTER-ALPHA-INHIBITOR (I-ALPHA-1) IS COMPOSED OF H1, H2 AND BIKUNIN. INTER-ALPHA-LIKE INHIBITOR (I-ALPHA-LI) OF H2 AND BIKUNIN. AND PRE-ALPHA-INHIBITOR (P-ALPHA-I) OF H3 AND BIKUNIN.  
-1- PPM: HEAVY CHAINS ARE INTERLINKED WITH BIKUNIN VIA A CHONDROITIN 4-SULFATE BRIDGE TO THEIR C-TERMINAL ASPARTATE.  
-1- SIMILARITY: BELONGS TO THE ITIH FAMILY.  
-1- SIMILARITY: CONTAINS 1 WMPA DOMAIN.  
-----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/announcements> or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
-----  
CC  
DR EMBL; X67055; CAA47439.1; -.  
DR EMBL; X14690; CAA32821.1; -.

DR Genew; HGNC:6168; ITIH3.  
DR MIM; 146650; -.  
DR InterPro; IPR002035; WVF\_A.  
DR Pfam; PF00092; Wva; 1.  
DR SMART; SM00327; WVA; 1.  
DR PROSITE; PSS0234; WVF; 1.  
KW Serine protease inhibitor; Repeat; Signal; Multigene family;  
KM Glycoprotein.  
FT SIGNAL 1 17 POTENTIAL.  
FT PROPEP 18 30  
FT CHAIN 31 647 INTER-ALPHA-TRYPSIN INHIBITOR HEAVY CHAIN H3.  
FT PROPEP 648 885  
FT DOMAIN 278 438 WVF.  
FT CARBOHYD 87 576 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT BINDING 647 576 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CONFLICT 344 344 CHONDROITIN 4-SULFATE, CROSS-LINK SITE.  
FT CONFLICT 357 357 R -> K (IN REF. 2).  
FT CONFLICT 846 846 N -> G (IN REF. 2).  
FT CONFLICT 846 846 Q -> H (IN REF. 2).  
SQ SEQUENCE 885 AA; 99121 MW; BC63856F8F5E4A1B CRC64;  
Query Match 8.2%; Score 92; DB 1; Length 885;  
Best Local Similarity 24.8%; Pred. No. 1.9;  
Matches 34; Conservative 30; Mismatches 45; Indels 28; Gaps 7;  
QY 16 YLVDSGSMKMKHVAVEPKIELAKAALKINAMPKSYGGLYTFAPYVLIPOGSMN 75  
DB 283 FVIDISGSMAGR-----KLEQTKELALRLLEMKEDY---LNFILPFGDV--STWK 329  
QY 76 SCVAECA-----VNTIKSDLEIFGRLLTPVGDGIKMETVIN-----QMPPOAA--VIL 121  
DB 330 EHLVQATPENLQEARTEFKVSKMEDKG-MTNINDGLRGISMINKAREHRIBERSTSIYM 388  
QY 122 LTGDHNNLGNMPVEVK 138  
DB 389 LTGDANVGESRPEKIQ 405  
RESULT 4  
ID TNFA MARMO STANDARD; PRT; 233 AA.  
AC 035734;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Tumor necrosis factor precursor (TNF-alpha) (Tumor necrosis factor ligand superfamily member 2) (Cachectin).  
GN TNF OR TNFSF2 OR TNFA.  
OS Mammota monax (Woodchuck).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Sciurinae;  
OC Mammota.  
OC NCBI\_Taxid=9995;  
OX NCBI\_Taxid=9995;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Peritoneal blood;  
RX MEDLINE=88139533; PubMed=9472070;  
RA Lorengel B., Lu M., Roggendorf M.;  
RT "Molecular cloning of the woodchuck cytokines: TNF-alpha, IFN-gamma, and IL-6.";  
RL Immunogenetics 47:332-335(1998).  
[2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Peritoneal blood;  
RX MEDLINE=20184748; PubMed=10721723;  
RA Li D.H., Haveli E.A., Brown C.L., Cullen J.M.;  
RT "Woodchuck lymphotoxin-alpha, -beta and tumor necrosis factor genes;

RT structure, characterization and biological activity.";  
RL Gene 242:295-305(2000).  
CC -1- FUNCTION: Cytokine that binds to TNFRSF1A/TNFR1 and  
CC TNFRSF1B/TNFR. It is mainly secreted by macrophages and can  
CC induce cell death of certain tumor cell lines. It is potent  
CC pyrogen causing fever by direct action or by stimulation of  
CC interleukin 1 secretion and is implicated in the induction of  
CC cachexia, under certain conditions it can stimulate cell  
CC proliferation and induce cell differentiation.  
CC -1- SUBUNIT: Homotrimer (By similarity).  
CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Also exists as an  
CC extracellular soluble form (By similarity).  
CC -1- PTM: The soluble form derives from the membrane form by  
CC proteolytic processing (By similarity).  
CC -1- PTM: The membrane form, but not the soluble form, is  
CC phosphorylated on serine residues. Dephosphorylation of the  
CC membrane form occurs by binding to soluble TNFRSF1A/TNFR1 (By  
CC similarity).  
CC -1- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING  
CC CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH  
CC AND MALNUTRITION.  
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; Y14137; CAA74569.1; -;  
DR EMBL; AF082491; AAC32615.1; -;  
DR EMBL; AF012910; AAF34863.1; -;  
DR HSSP; P06804; 2TNF.  
DR InterPro; IPR003636; TNF\_family.  
DR InterPro; IPR00478; TNF\_family.  
DR Pfam; PF00229; TNF\_1.  
DR PRINTS; PR01234; TNECROSISFCT.  
DR PRODOM; PD002012; TNF\_abc; 1.  
DR SMART; SM00207; TNF\_1.  
DR PROSITE; PS00251; TNF\_1; 1.  
DR PROSITE; PS00049; TNF\_2; 1.  
KW Cytokine; Cytotoxin; Transmembrane; Phosphorylation; Signal-anchor.  
FT CHAIN 1 233 TUMOR NECROSIS FACTOR, MEMBRANE FORM.  
FT DOMAIN 78 233 TUMOR NECROSIS FACTOR, SOLUBLE FORM.  
FT TRANSMEM 36 56 CYTOPLASMIC (POTENTIAL).  
FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
FT (POTENTIAL).  
FT DOMAIN 57 233 EXTRACELLULAR (POTENTIAL).  
FT SITE 77 78 CLEAVAGE (BY ADAM17) (BY SIMILARITY).  
FT MOD\_RES 2 2 PHOSPHORYLATION (BY CK1) (BY SIMILARITY).  
FT DISULFID 146 177 BY SIMILARITY.  
SQ SEQUENCE 233 AA; 25764 MW; 34D3D1965DAAD0E7 CRC64;  
  
Query Match 8.1%; Score 91; DB 1; Length 233;  
Best Local Similarity 23.1%; Pred. No. 0.44;  
Matches 45; Conservative 28; Mismatches 52; Indels 70; Gaps 10;

Db 142 KG0GCPSYLLTHTV 156  
RESULT 5  
ID DHE4\_EMENT STANDARD; PRT; 459 AA.  
AC P18619;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE NADP-specific glutamate dehydrogenase (EC 1.4.1.4) (NADP-GDH).  
GN GDH4.  
OS Emericella nidulans (Aspergillus nidulans).  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
OC Euryales; Trichocomaceae; Emericella.  
OX NCBI\_TaxId=5072;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89384423; PubMed=2550758;  
RA Hawkins A.R., Gur S.J., Montague P., Kinghorn J.R.;  
RT "Nucleotide sequence and regulation of expression of the Aspergillus  
RT nidulans gdhA gene encoding NADP dependent glutamate dehydrogenase";  
RL Mol. Gen. Genet. 218:105-111(1989).  
CC -1- CATALYTIC ACTIVITY: L-glutamate + H(2)O + NADP(+) = 2-oxoglutarate  
CC + NH(3) + NADPH.  
CC -1- SIMILARITY: BELONGS TO THE GLU/LEU/PHE/VAL DEHYDROGENASES FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X16121; CAA34252.1; -;  
DR PIR; S04904; S04904.  
DR HSSP; P24295; 1AUP.  
DR InterPro; IPR001625; GLFV\_Dh.  
DR Pfam; PF00208; GLFV\_dehydrog; 1.  
DR Pfam; PF02812; GLFV\_dehydrog\_N; 1.  
DR PRINTS; PR00082; GLFDHGRNASE.  
DR PROSITE; PS00074; GLFV\_DEHYDROGENASE; 1.  
KW Oxidoreductase; NADP.  
FT ACT\_SITE 114 114 BY SIMILARITY.  
SQ SEQUENCE 459 AA; 49608 MW; 682964399C002BD CRC64;  
  
Query Match 8.0%; Score 90; DB 1; Length 459;  
Best Local Similarity 24.2%; Pred. No. 1.3;  
Matches 45; Conservative 25; Mismatches 60; Indels 56; Gaps 8;

DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Glutamate racemase (EC 5.1.1.3)  
 GN MURI OR RV1338 OR MT1379 OR MTCY130.23 OR MTCY02B10.02.  
 OS Mycobacterium tuberculosis  
 OC Bacteria; Actinobacteriales; Actinobacteria (class); Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 NX NCBI\_TaxID=1773;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H37Rv.  
 RX MEDLINE=96295987; PubMed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eiglmeier K., Gae S., Barry C.E. III, Tekala F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
 RA Sultson J.E., Taylor K., Whitehead S., Barrett B.G.,  
 RT "Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence."  
 RL Nature 333:537-544(1998).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=CDC 1551 / Oshkosh;  
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,  
 RA Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,  
 RA Delcher A., Utterback T., Weidman J., Kouri H., Gill J., Mikula A.,  
 RA Bishai W.,  
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
 RT laboratory strains";  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: PROVIDES THE (R)-GLUTAMIC ACID REQUIRED FOR CELL  
 CC WALL BIOSYNTHESIS (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: L-glutamate = D-glutamate.  
 CC -1- PATHWAY: Peptidoglycan biosynthesis.  
 CC -1- SIMILARITY: BELONGS TO THE ASPARTATE/Glutamate RACEMASES FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Z73902; CA98102.1; -.  
 DR EMBL; AB007011; AAK45644.1; -.  
 DR HSSP; P56868; 1B74.  
 DR TIGR; MT1379; -.  
 DR Tuberculin; RV1338; -.  
 DR InterPro; IPR001920; Asp/Glu race.  
 DR InterPro; IPR004391; Glu race.  
 DR Pfam; PF01177; Asp\_Glu\_race; 1.  
 DR TIGRPFAMs; TIGR00067; Glu\_race; 1.  
 DR PROSITE; PS00923; ASP\_GLU\_RACEMASE\_1; 1.  
 DR PROSITE; PS00924; ASP\_GLU\_RACEMASE\_2; 1.  
 KW Peptidoglycan synthetis; Cell wall; Isomerase; Complete proteome.  
 SQ SEQUENCE 271 AA; 28643 MW; 29DE3CE67A27EEBA CRC64;

Query Match 7.5%; Score 84; DB 1; Length 271;

Best Local Similarity 24.8%; Pred. No. 2.3; Mismatches 19; Indels 18; Gaps 3;

QY 96 LTPVG-----DGKMHETVYNQMPPOAAILITDGHNNLTGMNPEVEKSIYQTNPNVCF 149

DB 5 LAPVGVPSGVGGLTVARAIQDLDEPDIVYVGDTONGVGFLTIPETIRA----- 54

QY 150 HVVSFADDAEGKAIIDQIVALNSGS--VLVDGLQLLQNPAY 188

DB 55 KALAIQDPLVGRGVKALVIACNSASSACLRDRRERYQVPV 95

RESULT 7

PGK\_RHILLO STANDARD; PRT; 398 AA.

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DE Phosphoglycerate kinase (EC 2.7.2.3).

GN PGK OR MUR3753.

OS Rhizobium loti (Mesorhizobium loti).

OC Bacteria; Proteobacteriales; alpha subdivision; Rhizobiaceae group;

OC Phyllobacteriaceae; Mesorhizobium.

NX NCBI\_TaxID=381;  
 [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MAFF303099;

RX MEDLINE=21082930; PubMed=11214968;

RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,

RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,

RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,

RA Wochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,

RA Takeuchi C., Yamada M., Tabata S.;

RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium

RT Mesorhizobium loti."

RL DNA Res. 7:331-338(2000).

CC -1- CATALYTIC ACTIVITY: ATP + 3-phospho-D-glycerate = ADP + 3-

CC phospho-D-glyceroyl phosphate.

CC -1- PATHWAY: Second phase of glycolysis; second step.

CC -1- SUBUNIT: Monomer (By similarity).

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potentially).

CC -1- SIMILARITY: BELONGS TO THE PHOSPHOGLYCERATE KINASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; AP003002; BAB50576.1; -.  
 DR InterPro; IPR001576; PGK.  
 DR Pfam; PF00162; PGK; 1.  
 DR PRINTS; PR00477; PHGLYCKINASE.  
 DR PROSITE; PS00111; PGLYCKINASE; 1.  
 KW Transferase; Kinase; Glycolysis; Complete proteome.  
 SQ SEQUENCE 398 AA; 41304 MW; 965669CB04C0A5FD CRC64;

Query Match 7.3%; Score 81.5; DB 1; Length 398;

Best Local Similarity 26.6%; Pred. No. 6.1; Mismatches 52; Indels 37; Gaps 6;

QY 86 IKSDEIFGRRLTPVGDGKMHETVYNQMPPO-----AAVILLITD-GHNNLTGMNPEVE 136

DB 20 VRVDLNV-----PVADGVNTPATRIERAIPTIABSGKAVIILLAHGRPRDGSPEFS 74

QY 137 VKSIQTNPNVCFHVSPADDAEGKAIIDQIVALNSGVLVDGLQLLQNPAY 185

DB 75 LEPIARATVATVIGRPVGRASDCVGMGSAVAAMNKGVL-----LFENTPFPYKAEKND 129

QY 186 PAVCOE-----FNSVIF 197

DB 130 PAFSERLANGDITFVNDAF 148

RESULT 8

YAVL\_XANCV STANDARD; PRT; 1164 AA.

ID YAVL\_XANCV STANDARD; PRT; 1164 AA.

DT 01-APR-1990 (Rel. 14, Created)



Db 331 DMLVATPANK-EAKTFVKIHDQSMNTINDGLKGIEMANKAREDTVEPSTIIM 389

QY 122 LTGDHNNLGNMVEEVK 138

Db 390 LTGDGANTGESRPEKIQ 406

RESULT 10

CADH\_MEDSA STANDARD; PRT; 358 AA.

AC P1656;

DT 01-JUN-1993 (Rel. 26, Created)

DT 01-JUN-1993 (Rel. 26, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Cinnamyl-alcohol dehydrogenase (EC 1.1.1.195) (CAD).

OS Medicago sativa (Alfalfa).

OC Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

OX NCBI\_TaxID=3879;

[1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Apollo;

RA van Doorselaere J., Baucher M., Feuillet C., Boudet A.M.,

RT "Isolation of cinnamyl alcohol dehydrogenase cDNAs from two important economic species: alfalfa and poplar. Demonstration of a high homology of the gene within angiosperms."

RL Plant Physiol. Biochem. 33:105-109(1995).

[2]

RP SEQUENCE FROM N.A., AND CHARACTERIZATION.

RC STRAIN=cv. Silver; TISSUE=Stem;

RX MEDLINE=20044095; PubMed=10579494;

RA Bill E.M., Abrahams S., Hayes C.M., Jenkins C.L., Watson J.M.;

RT "Molecular characterization and expression of a wound-inducible cDNA encoding a novel cinnamyl-alcohol dehydrogenase enzyme in lucerne (Medicago sativa L.)."

RL Plant Mol. Biol. 41:279-291(1999).

CC -1- FUNCTION: THIS PROTEIN CATALYZES THE FINAL STEP IN A BRANCH OF PHENYLPROPANOID SYNTHESIS SPECIFIC FOR PRODUCTION OF LIGNIN MONOMERS. IT ACTS ON CONFERTL-, SINAPYL-, 4-COUMARYL- AND CINNAMYL-ALCOHOL.

CC -1- CATALYTIC ACTIVITY: cinnamyl alcohol + NADP(+) = cinnamaldehyde + NADPH.

CC -1- COFACTOR: ZINC (BY SIMILARITY).

CC -1- PATHWAY: Lignin synthesis.

CC -1- TISSUE SPECIFICITY: MOST ACTIVELY EXPRESSED IN STEM, HYPOCOTYL AND ROOT TISSUE.

CC -1- SIMILARITY: BELONGS TO THE ZINC-CONTAINING ALCOHOL DEHYDROGENASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/announcements> or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

CC -----

DR EMBL; Z19573; CAAT9625.1; -.

DR EMBL; AF083332; AAC35845.1; -.

DR PIR; S31572; S31572.

DR InterPro: IPR002328; ADH\_zinc.

DR InterPro: IPR002085; Adh\_zn\_family.

DR Pfam; PF00107; adh\_zinc; 1.

DR PROSITE; PS00059; ADH\_ZINC; 1.

KW Oxidoreductase; NADP; Zinc; Lignin biosynthesis.

FT METAL 48 48 ZINC (CATALYTIC) (BY SIMILARITY).

FT METAL 70 70 ZINC (CATALYTIC) (BY SIMILARITY).

FT METAL 101 101 ZINC (SECOND ATOM) (BY SIMILARITY).

FT METAL 104 104 ZINC (SECOND ATOM) (BY SIMILARITY).

FT METAL 107 107 ZINC (SECOND ATOM) (BY SIMILARITY).

FT METAL 115 115 ZINC (SECOND ATOM) (BY SIMILARITY).

FT METAL 164 164 ZINC (CATALYTIC) (BY SIMILARITY).

SQ SEQUENCE 358 AA; 38948 MW; FBA609408D01BF56 CRC64;

Query Match 7.2%; Score 80.5; DB 1; Length 358;

Best Local Similarity 22.7%; Pred. No. 6.6;

Matches 63; Conservative 35; Mismatches 69; Indels 11; Gaps 15;

QY 1 AEVTASCTKRVESYNYLVDSGSMKGVAVREPKIELAKEIILKINAMPMKSYOGGLY 60

DB 109 SEIEQCNCKIKWISNDV--YIDG-----KIT-----QGG-- 135

QY 61 TPAPYSVI-----IPQGSNCSVAB--CAVNTIKSDLEIFGRLLP-----VG 100

DB 136 -FAESTVBEQKFFVVKIPEGLAPEQVAPLPCAGVTVPSPHSFGKTPGLRGIIIGAGVG 194

QY 101 D-GIKM-----HETVINQMPQAAVILLIDGHN-----LGM----- 131

DB 195 HMGVAKALGHVTVLSSDKKKKLELDYADNYLVSSDTVMQEAADSLDYITVP 254

QY 132 --NPVEEVKSIYQTNPNVCFHVSPADDAEGKAIIDQVALNSGSLVDGLQLQNPVAVC 189

DB 255 VGHPLPEYLSLK-----IDGKLIMGVI--NTPLOQFTPMVMIGRKSIT 297

QY 190 QEFVNSV-----FCOEQLVTE-EVYVLKGVNPAF 218

DB 298 GSFVGSVETEMLEFMWKEKGLTSMIRIYTDYINKAF 335

RESULT 11

GPDI\_YEAST STANDARD; PRT; 391 AA.

AC Q00055;

DT 01-JUN-1993 (Rel. 26, Created)

DT 01-JUN-1993 (Rel. 26, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Glycero1-3-phosphate dehydrogenase [NAD+] 1 (EC 1.1.1.1).

GN GPDI OR OSGI OR DARI OR HORI OR YDLO22W OR D2830.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.

OX NCBI\_TaxID=4932;

RP SEQUENCE FROM N.A.

RP MEDLINE=94254870; PubMed=8196651;

RA Alberlyn J., Hochmann S., Thevelein J.M., Prior B.A.;

RT "GPDI, which encodes glycerol-3-phosphate dehydrogenase, is essential for growth under osmotic stress in Saccharomyces cerevisiae, and its expression is regulated by the high-osmolality glycerol response pathway."

RL Mol. Cell. Biol. 14:4135-4144(1994).

[2]

RP SEQUENCE FROM N.A.

RX MEDLINE=94254870; PubMed=8196651;

RA Alberlyn J., Hochmann S., Thevelein J.M., Prior B.A.;

RT "GPDI, which encodes glycerol-3-phosphate dehydrogenase, is essential for growth under osmotic stress in Saccharomyces cerevisiae, and its expression is regulated by the high-osmolality glycerol response pathway."

RL Mol. Cell. Biol. 14:4135-4144(1994).

[3]

RP SEQUENCE FROM N.A.

RC STRAIN=J132B;

RX MEDLINE=95050285; PubMed=7961476;

RA Wang H.T., Rahaim P., Robbins P., Vocum R.R.;

RT "Cloning, sequence, and disruption of the Saccharomyces diastaticus DARI gene encoding a glycerol-3-phosphate dehydrogenase.";

RL J. Bacteriol. 176:7091-7095(1994).

[4]

RP SEQUENCE FROM N.A.

RC STRAIN=S288C;

RA Andre B., Ureastazu L.;

RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.

[5]

RP SEQUENCE OF 1-102 FROM N.A.

```

RX MEDLINE=91285439; PubMed=1676389;
RA Sleep D., O'Gden J.E., Roberts N.A., Goodey A.R.;
RT "Cloning and characterisation of the Saccharomyces cerevisiae
RL glycerol-3-phosphate dehydrogenase (GUT2) promoter.";
RN Gene 101:89-96(1991).
RP SEQUENCE OF 8-16; 95-108 AND 160-173.
RC STRAIN=ATCC 38531 / Y41;
RX MEDLINE=97089742; PubMed=8935550;
RA Norbeck J., Blomberg A.;
RT "Protein expression during exponential growth in 0.7 M NaCl medium of
RL Saccharomyces cerevisiae.";
RN FEMS Microbiol. Lett. 137:1-8(1996).
CC -1- CATALYTIC ACTIVITY: Sn-glycerol 3-phosphate + NAD(+) = glycerone
phosphate + NADH.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (potential).
CC -1- INDUCTION: BY OSMOTIC STRESS.
CC -1- SIMILARITY: BELONGS TO THE NAD-DEPENDENT GLYCEROL-3-PHOSPHATE
DEHYDROGENASE FAMILY.
CC -1- CAUTION: WAS ORIGINALLY (REF.5) THOUGHT TO BE GUT2, THE FAD-
DEPENDENT GLYCEROL-3-PHOSPHATE DEHYDROGENASE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z24454; CAA80827.1; -
DR EMBL; Z76859; CAA54189.1; -
DR EMBL; U04621; AAA64936.1; -
DR EMBL; Z48432; CAA88337.1; -
DR EMBL; Z74071; CAA98582.1; -
DR EMBL; M38740; AAA18631.1; -
DR PIR; P00148; P00148.
DR PIR; S34641; S34641.
DR PIR; S40059; S40059.
DR SGD; S0002180; GPD1.
DR InterPro; IPR001652; NAD-Gly3P-dh.
DR Pfam; PF01210; NAD_Gly3P_dh; 1.
DR PRINTS; PR00077; GPDHNRGNASE.
DR PRODOM; PD001649; NAD_Gly3P_dh; 1.
DR PROSITE; PS00957; NAD_G3PDH; 1.
DR Oxidoreductase; NAD; Multigene family.
KT CONFLICT 103 107 DNLVA -> TIVLL (IN REF. 5; AAA18631).
SQ SEQUENCE 391 AA; 42869 MW; F5F1F7F11F707D1 CRC64;

Query Match 7.1%; Score 80; DB 1; Length 391;
Best Local Similarity 22.1%; Pred. No. 8.1;
Matches 40; Conservative 27; Mismatches 58; Indels 56; Gaps 10;

OY 64 PYSV-IIPOGSWNSCVAEAVNTIKSDELIFGRITPVG-----DGIKHEIV-----I 110
DB 34 PPKVTVTIGSMNGTITAKVAENCKYPEVFAPIVQMWVEEINEKLEITINTHQNV 93
OY 111 NQMPPOAAVITLTDGNNNGAMP--VEEYKSI-----YQNPVNVCHVVSFADDAEG 160
DB 94 KYLIFG-----ITLFD--NLVANPDILDSYDVIDIVFNIPHQFLPRICQLKQHV----- 142
OY 161 KALIDIVALNNGSVLVDGILLONPAVCOEFNVSFCQGLIVTEVVV-----LGVNP 216
DB 143 -SHVRAISCKLGEFVNAKGVOLLSS-----YITTEIGICGALSGANI 184
OY 217 A 217
DB 185 A 185

RESULT 12
YA79_SCHPO STANDARD; PRT; 753 AA.
ID YA79_SCHPO

```

```

AC Q09763;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein C24Hc.09 in chromosome I.
GN SPAC24Hc.09.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
CX NCBI_TaxID=4896;
OX [1]
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brocks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,
RA Gentles S., Gobie A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Voicakeert G., Aert R., Robben J., Grympeprez B.,
RA Weljens J., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambut R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaue V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Roche M., Galliard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Ray F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cernutski L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrett B.G., Nure P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -1- SIMILARITY: CONTAINS 1 DBL-HOMOLOGY (DH) DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z54142; CAA90853.1; -
DR InterPro; IPR001331; GDS_CDC24.
DR InterPro; IPR000219; RHOGEF.
DR Pfam; PF00621; RHOGEF; 1.
DR SMART; SM00325; RHOGEF; 1.
DR PROSITE; PS00741; DH_1; FALSE_NEG.
DR PROSITE; PS0010; DH_2; 1.
KW Hypothetical protein.
KW DOMAIN 311 507 DH.
SQ SEQUENCE 753 AA; 84485 MW; 5740A1AED67F207 CRC64;

Query Match 7.1%; Score 80; DB 1; Length 753;
Best Local Similarity 21.5%; Pred. No. 18;
Matches 48; Conservative 38; Mismatches 85; Indels 52; Gaps 10;

OY 12 ESYNYLVDSYSGMMKHAIVAREPKIELAKEALIKINAAAPKNSYOGGLV-----TFAPY- 65
DB 268 ESLSRITISVATISLPSTPGKRLSKSLANSEASINPEYKRIKCANLIKELVTEAAVL 327
OY 66 -SVIIPQGSWNSCVAEC-AVNTIKSD-----LIFGRITPVGDG---IKHE 107
DB 328 NDILAIQDSYGLRVEKCSALNPVDAQTVGDIESTLITFTVEPHSRIRYQAGSGSMRVNLD 387

```





```

CC -----
DR EMBL: AJ010302; CAB38734.1; -.
DR EMBL: AF195122; AAF24284.1; -.
DR InterPro: IPR002667; IPR isomerase.
DR InterPro: IPR000086; NUDIX_hydrolase.
DR Pfam: PF00293; NUDIX.1.
DR ProDom: PD004109; IPR isomerase; 1.
DR Isomerase; Isoprene biosynthesis; Photosynthesis;
KW Chlorophyll biosynthesis.
FT ACT SITE 62 62 BY SIMILARITY.
FT ACT SITE 110 110 BY SIMILARITY.
SQ SEQUENCE 177 AA; 20295 MW; 84AB789DF17ADFAB CRC64;

Query Match 7.1%; Score 79; DB 1; Length 177;
Best Local Similarity 23.4%; Pred. No. 3.7;
Matches 30; Conservative 26; Mismatches 44; Indels 28; Gaps 6;

QY 26 MKHVAAREPKIELAKKINAAKPMKSYQGGLTFAPYSVIIPQGSNNSCVAECVANT 85
DB 26 LKHMAI--SVFWAGRAVLIGRRAGKY-HTPGLMA-----NTCCTPRWGEADCAVR 78
QY 86 IKSDEIFERLT-----PVGDGIMHETVINQMPQAAVILLTGHNHNGKMPVE 135
DB 79 LBEELGITLVTFADREYRADVGNGLTEHEVV-----DIFVAAPSPDLVNP-- 127
QY 136 EVKSIYQT 143
DB 128 DPEWET 135

RESULT 15
THRC_HELPY STANDARD; PRT; 486 AA.
AC O24824:
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Threonine synthase (EC 4.2.3.1).
GN THRC OR HP0098.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCB1_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26585 / ATCC 700392;
RK MEDLINE=9739467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khaliq H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Wathey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RA "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori."
RL Nature 388:539-547(1997).
CC -1- CATALYTIC ACTIVITY: O-phospho-L-homoserine + H(2)O = L-threonine +
CC phosphate.
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC -1- PATHWAY: Threonine biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE SERINE/THREONINE DEHYDRATASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

DR EMBL: AE000531; AAD07166.1; -.
DR TIGR: HP0098; -.
DR InterPro: IPR001926; B6 enzyme Beta.
DR InterPro: IPR000634; S/T dehydrase.
DR InterPro: IPR004450; Thr_synthase.
DR Pfam: PF00291; PALP.1.
DR TIGRFAMs: TIGR00260; thrc.1.
DR PROSITE: PS00165; DEHYDRATASE_SER_THR; 1.
KW Threonine biosynthesis; Lyase; Pyridoxal phosphate; Complete proteome.
FT BINDING 109 109 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 486 AA; 54706 MW; CD9637799B508988 CRC64;

Query Match 7.1%; Score 79; DB 1; Length 486;
Best Local Similarity 21.8%; Pred. No. 13;
Matches 41; Conservative 28; Mismatches 73; Indels 46; Gaps 10;

QY 30 AVREPKIELAKKINAAKPMKSYQGGLTFAPYSVIIPQGSNNSC-----VAECA 82
DB 8 SLKEKKIDFT-EALINPNA--PK---GGLYTLERFETL-----QWQCLNLSYNDIVECV 56
QY 83 VNTIKSDEIFGRLTVPVGDGIMHETVINQMPQAAVILLTD-----GHNHL 129
DB 57 PERL--GLEIPKNL--LASALKRYENFDPNPKNP-APFALNERLFGVQLYHGPSLAFKDM 111
QY 130 GNPVEEVKSIYQINPNVCFHVSPADDAEGKAITDQVAL-----NSGSVLYDG 179
DB 112 ALQPLASLFSNLAVGKNEKYLMLVSTSGDTGPATLESLAGMPNPFVCLYPRDGTSLVQK 171
QY 180 LQLLNPA 187
DB 172 LQMTQSA 179

```

Search completed: April 9, 2003, 14:30:39  
 Job time : 29 secs



GenCore version 5.1.4 p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 9, 2003, 14:28:37 ; Search time 44 Seconds  
(without alignments)  
476.302 Million cell updates/sec

Title: US-10-034-500-2

Perfect score: 1119

Sequence: 1 AEVTASCTKRVESYNYLVY.....QEQILVTEVVLGVNFAF 218

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: p1r73:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	126	11.3	227	1 G70062	conserved hypochet
2	106.5	9.5	224	1 H70062	conserved hypochet
3	99	8.8	426	2 A96268	phosphoglycerate k
4	97.5	8.7	400	2 AG3016	phosphoglycerate k
5	93.5	8.4	2848	2 T32550	hypothetical prote
6	92.5	8.3	615	2 AH2028	hypothetical prote
7	92	8.2	885	2 S30350	inter-alpha-trypsi
8	90.5	8.1	347	2 G82862	conjugal transfer
9	90	8.0	318	2 E82493	conserved hypochet
10	90	8.0	351	2 S52708	hypothetical prote
11	90	8.0	459	1 S04904	glutamate dehydrog
12	90	8.0	903	2 G82516	hypothetical prote
13	90	8.0	1091	2 AF1380	glycosidase homolo
14	87.5	7.8	932	2 UC5853	inter-alpha-inhibi
15	86	7.7	1090	2 AG1749	glycosidase homolo
16	84	7.5	509	2 T12521	probable glutamate
17	83.5	7.5	271	2 T12521	hypothetical prote
18	83	7.4	1313	2 G82887	hypothetical prote
19	82.5	7.4	209	2 B82501	hypothetical prote
20	82	7.3	913	2 TG0168	gob-5 protein - mo
21	82	7.3	2205	2 T08615	aggregation factor
22	81.5	7.3	308	2 E71697	probable proteinas
23	81.5	7.2	1187	2 T45635	hypothetical 125K
24	81	7.2	547	2 T45635	hypothetical prote
25	81	7.2	886	2 S54355	inter-alpha-trypsi
26	80.5	7.2	339	2 E82211	conserved hypochet
27	80.5	7.2	358	2 S31572	cinamyl-alcohol d
28	80	7.1	110	2 S36258	lg lambda chain V
29	80	7.1	391	2 S40059	glycerol-3-phospha

30	80	7.1	476	2 T40086	hypothetical prote
31	80	7.1	753	2 S62411	probable guanine n
32	80	7.1	2055	2 T31110	extracellular matr
33	79.5	7.1	1165	2 T38867	probable DNA-dirc
34	79	7.1	177	2 T50740	isopentenyl diphos
35	79	7.1	357	2 AH1893	hydrogenase expres
36	79	7.1	468	2 S61964	probable membrane
37	79	7.1	486	2 B64532	threonine synthase
38	79	7.1	601	2 C81715	gcpE protein TC032
39	79	7.1	670	2 A12223	transketolase (imp
40	79	7.1	723	2 PN0509	integrin beta-3 ch
41	79	7.1	875	1 ITECAP	DNA topoisomerase
42	79	7.1	875	2 B91018	DNA gyrase subunit
43	79	7.1	875	2 D85862	DNA gyrase, subunit
44	79	7.1	1652	2 I50711	complement C3 prec
45	79	7.1	2183	1 G48556	genome polypeptide

#### ALIGNMENTS

##### RESULT 1

G70062 conserved hypothetical protein ywmc - Bacillus subtilis

C/Species: Bacillus subtilis

C/Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000

C/Accession: G70062

R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch

A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gal

lechi, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holzapfel, S.; Hosono, S.; Hullo, M.F

Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois

A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue

Y, M.; Ogawa, K.; Ogihara, A.; Oudeg, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetell

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon

A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sero

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpestra, P.; Tognoni, A.; Tosato, V.; Uchiyama

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,

A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Zumbstein, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A:Reference number: A69580; PMID:98044033; PMID:9384377

A:Accession: G70062

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-227 <KUN>

A:Cross-references: GB:299122; GB:AL009126; NID:G2636029; PIDN:CAB15691.1; PID:G2636199

A:Experimental source: strain 168

C:Gene: ywmc

C:Superfamily: hypothetical protein ywmc

Query Match 11.3%; Score 126; DB 1; Length 227;

Best local similarity 24.6%; Pred. No. 0.0014;

Matches 57; Conservative 37; Mismatches 84; Indels 54; Gaps 11;

QY	1	AEVTASCTKRVESYNYLVYSGMMKGVAVREKIELAKVITKINAMP-----KMSY 55
DB	24	AETKTEPAKAPANAVALLDASGSMARKIDGV--SKFNSAKKEISKFSASPREGQVKNVSV 81
QY	56	QG-----GLTFEAPYSVILIPGGSNSCAEAVNVTISDLIFPGLT 97
DB	82	FGSEGNKNSGKQSCAIVNRYGFSFN--EDSFNS-----LNTIGT-----GWT 127
QY	98	PVGDGIRHETVINQMP--QAAVILLTDGHNINGNPVEVKSIVOTNPVCHVVSFA 155
DB	128	PIAKALNEAKSPFDQDAKEKVVYLLTDGERTCGGPIKAKELQK--DNITNVVIGF- 184
QY	156	DDAEG-KAIIIDQIVAAVNSGVLDGLOLQNPVAVCGFVNSVFCQEQILVTE 206
DB	185	DYKGGYGDALNAIAKVGGEYF-----PAYTQDVERKIFQDSIMLSK 227

## RESULT 2

H70062

conserved hypothetical protein ymmd - *Bacillus subtilis*C/Species: *Bacillus subtilis*

C/Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000

C/Accession: H70062

R/Kunze, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bacter

C./Brom, S.; Broillett, S.; Brusch, C.V.; Caldwell, B.; Capiano, V.; Carter, N.M.; Chd

A./Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A/Authors: Foulger, C.R.; Fujita, C.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier

lechi, J.; Harwood, C.R.; Hentut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.

Koetter, P.; Koningsstein, G.; Krogsh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

A/Authors: lauber, J.; Lazarovic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel

Y, M.; Ogawa, K.; Ogilawa, A.; Oudaga, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,

A/Authors: Schleich, S.; Schroefer, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serot

akenchi, M.; Tamakoshi, A.; Tanaka, T.; Terpestra, P.; Tognoni, A.; Tosato, V.; Uchiiyama,

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K

A/Authors: Yoshikawa, H.F.; Zumelein, E.; Yoshikawa, H.; Danchin, A.

A/Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.

A/Reference number: AB5580; MUID:98044033; PMID:9384377

A/Accession: H70062

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-224 &lt;KUN&gt;

A/Cross-references: GB:299122; GB:AL009126; NID:92636029; PIDN:CAB15690.1; PID:g2636198

A/Experimental source: strain 168

C/Genetics:

A/Gene: ymmd

C/Superfamily: hypothetical protein ymmd

Query Match 9.5%; Score 106.5; DB 1; Length 224;

Best Local Similarity 21.9%; Pred. No. 0.067; Mismatches 66; Indels 55; Gaps 9;

Matches 42; Conservative 29; Mismatches 66; Indels 55; Gaps 9;

QY 17 LVYDSSMMKVAAREPEIKELAKELINAMPK-----MSYQGG----- 58

DB 35 LFDGSGSNVQK--TGERKIDIAKSSVKSFAELPKDNTMLMRVPHAGNNKLSGKALSC 92

QY 59 ----LYTFAPY-----SVIIPOGWSNCAEACAVNTKSDIEIFGRITPVGDGIK 104

DB 93 STTEITYGHPYEGSLFDNSLSLKPFG--WTP--IAKALADT--RKEFEAPD-----ADG-- 142

QY 105 MHEVTINQPPQAAVITLDGHNNGMNPVEYKSIYQNPVNCFFVSPFADAEKALII 164

DB 143 -----KNNVYLTIDGERTCGDPALETETKURASNVDTIVNIIGFEDVAKNEEM 191

QY 165 DQIVALLNSGSVL 176

DB 192 KQAAVAGGSEYI 203

## RESULT 3

A96268

phosphoglycerate kinase, pgk (AF256214) [imported] - *Agrobacterium tumefaciens* (strain CC/Species: *Agrobacterium tumefaciens*

C/Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 11-Jan-2002

C/Accession: A96268

R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,

A./Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.;

Science 294, 2323-2328, 2001

A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tum*

A/Reference number: A97359; PMID:11743194

A/Accession: A96268

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-426 &lt;KUN&gt;

A/Cross-references: GB:AB007870; PIDN:AAK89667.1; PID:g15159569; GSPDB:GN00170

C/Genetics:

A/Gene: AGR\_L\_2193

A/Map position: linear chromosome

C/Superfamily: phosphoglycerate kinase

Query Match 8.8%; Score 99; DB 2; Length 426;

Best Local Similarity 27.2%; Pred. No. 0.69; Mismatches 63; Indels 42; Gaps 7;

Matches 46; Conservative 18; Mismatches 63; Indels 42; Gaps 7;

QY 65 YSVIIPGWSNCAEACAVNTKSDIEIFGRIL-----TPVGDGIKMETVINQMPQ 116

DB 12 HSMVPELATIADQDRMPAPFKTIDLDIAGRVVVRVDLNPVADGKVTDAIRIETVAPT 71

QY 117 -----AAVITLDGHNNGMNPVEE--VKSIYQNPVNCFFVSPFADAEKALIIDQ 166

DB 72 ILEISSKCAKVIILAHFGRPKG--EPVAMSSQIVPVEDVLDHATISFATDCIGAPADA 130

QY 167 IVALNSGSVLVDGQLQION-----PAVCOE-----FVNSVF 197

DB 131 VAKMNDGIL-----LLENTFRHKEERKNDAPFVEELANGDIYVNDAF 174

## RESULT 4

AG3016

phosphoglycerate kinase pgk [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont)C/Species: *Agrobacterium tumefaciens*

C/Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 01-Feb-2002

C/Accession: AG3016

R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; Mccliel

Science 294, 2317-2323, 2001

A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Kreepan, W.; Perry, M.; Gordon-Kamm,

eter, E.W.

A/Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.

A/Reference number: AB2577; PMID:11743193

A/Accession: AG3016

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-400 &lt;KUN&gt;

A/Cross-references: GB:AB008689; PIDN:AAU44549.1; PID:g17742163; GSPDB:GN00187

A/Experimental source: strain C58 (Dupont)

C/Genetics:

A/Map position: linear chromosome

C/Superfamily: phosphoglycerate kinase

Query Match 8.7%; Score 97.5; DB 2; Length 400;

Best Local Similarity 28.6%; Pred. No. 0.86; Mismatches 47; Indels 39; Gaps 7;

Matches 40; Conservative 14; Mismatches 47; Indels 39; Gaps 7;

QY 86 IKSDEIFGRILTPVGDGIKMETVINQMPQ-----AAVITLDGHNNGMNPVEE- 136

DB 20 VAVDLNV-----PVADGKVTDAIRIETVAPTILESSKCAKVIILAHFGRPKG--EPVAM 73

QY 137 -VKSIYQNPVNCFFVSPFADAEKALIIDQIVALLNSGSVLVDGQLQION----- 185

DB 74 SLISQIVPVEDVLDHATISFATDCIGAPADAVAKMNDGIL-----LLENTFRHKEERK 128

QY 186 -PAVCOE-----FVNSVF 197

DB 129 DPVVEELANGDIYVNDAF 148

## RESULT 5

T32550

hypothetical protein F33D4.2a - *Caenorhabditis elegans*C/Species: *Caenorhabditis elegans*

C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 21-Jan-2000

C/Accession: T32550

R/Johnson, D.; Stellyes, L.

submitted to the EMBL Data Library, December 1997

A/Description: The sequence of *C. elegans* cosmid F33D4.

A/Reference number: Z21190

A/Accession: T32550

A/Status: preliminary; translated from GB/EMBL/DBJ



A:Map position: 3p13-3p12  
 C:Superfamily: inter-alpha-tryptsin inhibitor complex component II  
 C:Keywords: chondroitin sulfate proteoglycan; glycoprotein; heterodimer; proteinase inh;  
 F:1-19/Domain: signal sequence #status predicted <SIG>  
 F:20-29/Domain: propeptide #status predicted <PRO>  
 F:30-64/Product: inter-alpha-tryptsin inhibitor heavy chain 3 #status experimental <MAT>  
 F:648-885/Domain: carboxyl-terminal propeptide #status predicted <CTP>  
 F:87-576/Binding site: carboxyl-terminal (Asn) (covalent) #status predicted  
 F:647/Modified site: chondroitin sulfate ester carboxyl end (Asp) (in mature form) #stat

Query Match 8.2%; Score 92; DB 2; Length 885;

Best Local Similarity 24.8%; Pred. No. 7.2;

Matches 34; Conservative 30; Mismatches 45; Indels 28; Gaps 7;

OY 16 YLVDSGSMKGVAVREPKIELAKKINAAAPKMSYGGGLTFAPYSVITIPGSMN 75  
 DB 23 FVIDISGSMAGR-----KLEOTKEALIRILEDMOEY---LNFILFSGDV--STWK 329  
 OY 76 SCVAECA-----VNTIKSDLEIFGLTPVGDGIKMHETVIN-----CMPQAA--VIL 121  
 DB 330 EHLVQATPENLQEAKTFFVSMEDKG-MTNINDGLRGISMLKAREHRIPEKSTIVIM 388  
 OY 122 LTGDGNNIGAMPVEEYK 138  
 DB 389 LTGDGANVGESRPERKIQ 405

## RESULT 8

CG82862 conugal transfer protein XFa0015 [imported] - Xylella fastidiosa (strain 9a5c)

C/Species: Xylella fastidiosa

C/Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Sep-2000

C/Accession: G82862

R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen

Nature 406, 151-157, 2000

A>Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A/Reference number: A82515; MUID:20365717; PMID:10910347

A/Note: for a complete list of authors see reference number A59328 below

A/Accession: G82862

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-347 <SIN>

A/Cross-references: GB:A8003851; NID:g9112238; PIDN:AAF85584.1; GSPDB:GN00130; XFSC:XFAC

A/Experimental source: strain 9a5c

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carrero, D.M.; Carrer, H

as-Neto, E.; Docena, C.; El-Porty, H.; Facincani, A.F.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A/Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm

J.D.; Junqueira, M.L.; Kemper, E.L.; Kltajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig

chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, E

A/Authors: Martins, E.M.F.; Matukuma, A.Y.; Menck, C.F.M.; Miranca, E.C.; Miyaki, C.Y.

, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de sa, R.G.; Santeili, R.V.; Sawasak

A/Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv

M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

A/Reference number: A59328

A/Contents: annotation

C/Genetics:

A/Gene: XFa0015

A/Genome: plasmid

A/Note: plasmid pXF5.1

C:Superfamily: tumor-inducing plasmid pTic58 virB1 protein

Query Match 8.1%; Score 90.5; DB 2; Length 347;  
 Best Local Similarity 22.1%; Pred. No. 2.9;  
 Matches 47; Conservative 35; Mismatches 80; Indels 51; Gaps 9;

OY 18 VDSGSMKGVAVREPKIELAKKINAAAPKMSYGGGLTFAPYSVITIPGSMN 77  
 DB 144 VASGTLHYEVEL---LALKSNRIK--EFLKIAMQGL---TLVVGKTSKGT 192  
 OY 78 VABCAVNTIKSDLEIFGLTPVGDGIKMHETVINQMPQAAVILTDGNNIGAMPVEEY 137

DB 193 IGSITNCIPD-----ERLTVED---VHEFMILHHPKVKHLYFSRDEGSLINPKQAI 245  
 OY 138 KSIYQTNVNCVHVSFPDDA-----EGKAI--IDQIVANSGSLV 177  
 DB 246 ASCLMKRDRLLTBMRCDEMEFVKAVGSGHPGISTHAGALPAFQIVALLKDS--A 303  
 OY 178 DGLQLQNPVAVCOEFVNSVFCOEQLVTEVVV 210  
 DB 304 TGAHL-----DAVYIKKRVETVDIVL 325

## RESULT 9

CG82493

conserved hypothetical protein VCA0172 [imported] - Vibrio cholerae (strain N1661 sero

C/Species: Vibrio cholerae

C/Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001

C/Accession: B82493

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.

charlson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers,

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A/Reference number: A82035; MUID:20406833; PMID:10952301

A/Accession: B82493

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-318 <HET>

A/Cross-references: GB:A8004357; GB:A8003853; NID:g9657547; PIDN:AAF96085.1; GSPDB:GN00

A/Experimental source: serogroup O1; strain N1661; biotype El Tor

C/Genetics:

A/Gene: VCA0172

A/Map position: 2

Query Match 8.0%; Score 90; DB 2; Length 318;  
 Best Local Similarity 26.7%; Pred. No. 2.9;  
 Matches 36; Conservative 22; Mismatches 53; Indels 24; Gaps 6;

OY 58 GLTFAPYSVI-IPGSMNSVCAECVAVNTIKSDLEIFGLTPVGDGIKMHETVINQMP 115  
 DB 131 GLTFADHAYLQPLTDRQVYANQINQTV---LKLIGQTAIGSGIGLATITFIDSDAP 187  
 OY 116 QAAVILTDGNNIG-MNPVEEKSIVQTNPNV-----CFHVSFPDDA 158  
 DB 188 QVWILSLDSGNTAGVLPDLKANAKQVNTIYVGVAGEMVVDPLFSKXVTAQDL 247  
 OY 159 EGKAIIDQIVANSG 173  
 DB 248 DEKTL--QTATTTG 260

## RESULT 10

SS2708

hypothetical protein YDR202c - yeast (Saccharomyces cerevisiae)

N/Alternate names: hypothetical protein YD9346.12c

C/Species: Saccharomyces cerevisiae

C/Date: 19-May-1995 #sequence\_revision 01-Sep-1995 #text\_change 19-Apr-2002

C/Accession: SS2708

R:Oliver, K.; Harris, D.

submitted to the EMBL Data Library, March 1995

A/Reference number: S52697

A/Accession: SS2708

A/Molecule type: DNA

A/Residues: 1-351 <OLI>

A/Cross-references: EMBL:Z48784; NID:g755782; PID:g755794; GSPDB:GN00004; MIPS:YDR202C

A/Experimental source: strain AB972

C/Genetics:

A/Gene: SGD:RAV2; MIPS:YDR202C

A/Cross-references: SGD:S0002610

A/Map position: 4R

C:Superfamily: Saccharomyces cerevisiae hypothetical protein YDR202c

Query Match 8.0%; Score 90; DB 2; Length 351;



A:Experimental source: strain EGD-e  
C:Genetics:  
A:Gene: lmo2446

Query Match 8.0%; Score 90; DB 2; Length 1091;  
Best Local Similarity 22.0%; Pred. No. 14;  
Matches 56; Conservative 37; Mismatches 74; Indels 88; Gaps 15;

QY 21 SGSMKMHVAVREPKIELAKEALIKINAMPKMSYQ-GGLYTFAPY-----SVIIPQ 71  
DB 743 NGKSMRQAMADYPE-----DYNARDLDEQYMFQDGLVAPIVOEGOTEKEVYLPE 793  
QY 72 GSMNSCVAECAVNTIKSDLEIF-GRLLTPVGDGIKMHEVINQMP--PQAAVIL--LTDG 125  
DB 794 GEM-----VDINMGVHPGGETISYADV-DLLPVFAKAGALIIPMMMTDG 837  
QY 126 HNNLGMNPEVEEKSIYQTPNVCFHV-----VSFADAEKALID-----KAI 163  
DB 838 Y-QLQGQNVNDLKSY-----DNLTFRVYPSGNSSEYFDDVNGEMKDISVSEDFANEKYS 892  
QY 164 ID-----QIVALNSGSLVVDGLQLLQ-----NPVCOEFVNSV-----FCQEQI 202  
DB 893 VDLPMADETTMQVFTSTPTSTVTDGADYAKADTLDAFNEATGTYTDTVQNLTYIKTAA 952  
QY 203 LVTBEVVVLGVNFA 217  
DB 953 KDQAQAIVLGVNHA 967

## RESULT 14

JCS953  
Inter-alpha-inhibitor H4p heavy chain - rat  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 05-Feb-1999 #sequence\_revision 05-Feb-1999 #text\_change 21-Jul-2000  
C:Accession: JCS953  
R:Soury, E.; Olivier, E.; Daveau, M.; Hiron, M.; Claeysens, S.; Ristler, J.L.; Salier, J.  
Biochem. Biophys. Res. Commun. 243, 522-530, 1998  
A:Title: The H4p heavy chain of inter-alpha-inhibitor family largely differs in the stru  
A:Reference number: JCS953; MUID:98153798; PMID:9480842  
A:Accession: JCS953  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-932 <SOI>  
A:Cross-references: GB:Y11283; NID:92292987; PIDN:CAA72155.1; PID:92292988  
C:Superfamily: Inter-alpha-trypsin inhibitor complex component II

Query Match 7.8%; Score 87.5; DB 2; Length 932;

Best Local Similarity 22.1%; Pred. No. 19;  
Matches 33; Conservative 31; Mismatches 40; Indels 45; Gaps 9;

QY 16 YLVDYSGSMKMHVAVREPKIELAKEALIKINAMPKMSYQGLYTFAPYSVILIPQG--- 72  
DB 276 FYIDKSGSNAGK-----KIQOTREALIKI--LKDLSIOTD-----QENITVFSGEAN 319  
QY 73 SWNSCVAECAVNTIKSDLEIFGRLLTPVGDGIKMHETVI-----NQ---MPQAA-- 118  
DB 320 QMEQLIVQATEENINRAVDYASKI-PAQGTININKAVLSAVELLDRSNQALLPSKVSIL 378  
QY 119 VILLTDGHNHNLGMNPEVEEKSIYQTPNVC 147  
DB 379 IILLTDGEPTVG-----ETNPKI 396

## RESULT 15

AG1749

glycosidase homolog lin2540 [imported] - Listeria innocua (strain C1ip11262)  
C:Species: Listeria innocua  
C>Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AG1749  
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Biochecker  
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duesurget, O.; Entian, K.D.; Fsihi, H.  
Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madeno, E.; Maitournam, A.; M  
ok, C.; Schlueter, T.; Simoes, N.; Tietze, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland  
A:Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11619669

A:Accession: AG1749  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1090 <GLA>  
A:Cross-references: GB:AL592022; PIDN:CAC97767.1; PID:916415062; GSPDB:GN00178  
A:Experimental source: strain C1ip11262  
C:Genetics:  
A:Gene: lin2540

Query Match 7.7%; Score 86; DB 2; Length 1090;  
Best Local Similarity 21.6%; Pred. No. 31;  
Matches 55; Conservative 36; Mismatches 76; Indels 88; Gaps 14;

QY 21 SGSMKMHVAVREPKIELAKEALIKINAMPKMSYQ-GGLYTFAPY-----SVIIPQ 71  
DB 743 NGKSMRQAMADYPE-----DINTRNLDEQYMFQDGLVAPIVOEGOTEKEVYLPE 793  
QY 72 GSMNSCVAECAVNTIKSDLEIF-GRLLTPVGDGIKMHEVINQMP--PQAAVIL--LTDG 125  
DB 794 GEM-----VDINMGVHPGGETISYADV-DLLPVFAKAGALIIPMMMTDG 837  
QY 126 HNNLGMNPEVEEKSIYQTPNVCFHV-----VSFADAEKALID-----KAI 165  
DB 838 Y-QLQGQNVNDLKAY-----DNLTFRVYPSGNSSEYFDDVNGEMKDISVSEDFANEKYS 892  
QY 166 -----QIVALNSGSLVVDGLQLLQ-----NPVCOEFVNSV-----FCQEQI 202  
DB 893 VNLPMADETTMQVFTSTPTSTVTVAGAEVAKADTLDAFNEATSAVYDTVQNLTYIKTAA 952  
QY 203 LVTBEVVVLGVNFA 217  
DB 953 TDTTQAIVLGVNHA 967

Search completed: April 9, 2003, 14:33:04  
Job time : 48 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: April 9, 2003, 14:32:23 ; Search time 35 Seconds  
(without alignments)  
380.790 Million cell updates/sec

Title: US-10-034-500-2

Perfect score: 1119

Sequence: 1 AETVASCYKRVESYNYLVY.....QEQILVTEEVVLRGVNFAR 218

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AA.\*

1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB\_pdp.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PTCT\_NEW\_PUB\_pdp.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB\_pdp.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB\_pdp.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB\_pdp.\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB\_pdp.\*  
7: /cgn2\_6/ptodata/2/pubpaa/PTCTUS\_PUBCOMB\_pdp.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB\_pdp.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB\_pdp.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB\_pdp.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB\_pdp.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB\_pdp.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB\_pdp.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB\_pdp.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	92	8.2	885	10 US-09-828-423-5	Sequence 5, App1
2	91.5	8.2	917	9 US-09-981-353-54	Sequence 54, App1
3	91.5	8.2	917	12 US-10-025-167-41	Sequence 41, App1
4	91.5	8.2	919	9 US-09-992-598-379	Sequence 379, App
5	91.5	8.2	919	9 US-09-989-293A-379	Sequence 379, App
6	91.5	8.2	919	9 US-10-063-547-70	Sequence 70, App1
7	91.5	8.2	919	9 US-09-989-735-379	Sequence 379, App
8	91.5	8.2	919	9 US-09-990-444-379	Sequence 379, App
9	91.5	8.2	919	9 US-09-989-730-379	Sequence 379, App
10	91.5	8.2	919	9 US-09-990-436-379	Sequence 379, App
11	91.5	8.2	919	9 US-09-991-181-379	Sequence 379, App
12	91.5	8.2	919	9 US-09-993-687-379	Sequence 379, App
13	91.5	8.2	919	9 US-09-989-734-379	Sequence 379, App
14	91.5	8.2	919	9 US-09-997-653-379	Sequence 379, App
15	91.5	8.2	919	9 US-10-174-590-258	Sequence 258, App
16	91.5	8.2	919	9 US-10-176-758-258	Sequence 258, App
17	91.5	8.2	919	9 US-10-063-616-70	Sequence 70, App1
18	91.5	8.2	919	9 US-10-175-737-258	Sequence 258, App
19	91.5	8.2	919	9 US-09-993-667-379	Sequence 379, App

20	91.5	8.2	919	9 US-10-063-502-70	Sequence 70, App1
21	91.5	8.2	919	9 US-10-173-706-258	Sequence 258, App
22	91.5	8.2	919	9 US-10-175-738-258	Sequence 258, App
23	91.5	8.2	919	9 US-10-175-752-258	Sequence 258, App
24	91.5	8.2	919	9 US-10-176-482-258	Sequence 258, App
25	91.5	8.2	919	9 US-10-176-757-258	Sequence 258, App
26	91.5	8.2	919	9 US-10-176-913-258	Sequence 258, App
27	91.5	8.2	919	9 US-10-180-552-258	Sequence 258, App
28	91.5	8.2	919	9 US-10-180-557-258	Sequence 258, App
29	91.5	8.2	919	9 US-09-990-438-379	Sequence 379, App
30	91.5	8.2	919	9 US-09-990-562-379	Sequence 379, App
31	91.5	8.2	919	9 US-09-997-428-379	Sequence 379, App
32	91.5	8.2	919	9 US-09-997-666-379	Sequence 379, App
33	91.5	8.2	919	9 US-10-173-700-258	Sequence 258, App
34	91.5	8.2	919	9 US-10-174-572-258	Sequence 258, App
35	91.5	8.2	919	9 US-10-174-579-258	Sequence 258, App
36	91.5	8.2	919	9 US-10-174-582-258	Sequence 258, App
37	91.5	8.2	919	9 US-10-174-588-258	Sequence 258, App
38	91.5	8.2	919	9 US-10-175-739-258	Sequence 258, App
39	91.5	8.2	919	9 US-10-175-740-258	Sequence 258, App
40	91.5	8.2	919	9 US-10-175-743-258	Sequence 258, App
41	91.5	8.2	919	9 US-10-176-488-258	Sequence 258, App
42	91.5	8.2	919	9 US-10-176-492-258	Sequence 258, App
43	91.5	8.2	919	9 US-10-176-747-258	Sequence 258, App
44	91.5	8.2	919	9 US-10-176-750-258	Sequence 258, App
45	91.5	8.2	919	9 US-10-176-985-258	Sequence 258, App

## ALIGNMENTS

RESULT 1  
US-09-828-423-5  
Sequence 5, Application US/09828423  
Patent No. US20020099178A1

GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
Guegler, Karl J.  
Patterson, Chandra

TITLE OF INVENTION: GROWTH-ASSOCIATED TRYPsin-TYPE  
INHIBITOR HEAVY CHAIN PRECURSOR

NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESS: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 6.1/MS-DOS 6.2

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/828,423  
FILING DATE: 05-Apr-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/388,774  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Cerrone, Michael C  
REGISTRATION NUMBER: 39,132  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-845-0555  
TELEFAX: 650-845-4166  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 885 amino acids  
TYPE: amino acid







RESULT 4  
US-09-992-598-379  
Sequence 379, Application US/09992598  
Patent No. US20020160384A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi J.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Baton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gertltsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Guirney, Austin L.  
APPLICANT: Kijavini, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
Acids Encoding the Same  
FILE REFERENCE: P2730P1C20  
CURRENT APPLICATION NUMBER: US/09/992,598  
PRIOR FILING DATE: 2001-11-14  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030

PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23

PRIOR APPLICATION NUMBER: 60/090355  
 PRIOR FILING DATE: 1998-06-23  
 PRIOR APPLICATION NUMBER: 60/090429  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090431  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090435  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090444  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090445  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090472  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090535  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090540  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090542  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090557  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090676  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090678  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090690  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090694  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090695  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090696  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090862  
 PRIOR FILING DATE: 1998-06-26  
 PRIOR APPLICATION NUMBER: 60/090863  
 PRIOR FILING DATE: 1998-06-26  
 PRIOR APPLICATION NUMBER: 60/091560  
 PRIOR FILING DATE: 1998-07-01  
 PRIOR APPLICATION NUMBER: 60/091478  
 PRIOR FILING DATE: 1998-07-02  
 PRIOR APPLICATION NUMBER: 60/091544  
 PRIOR FILING DATE: 1998-07-01  
 PRIOR APPLICATION NUMBER: 60/091519  
 PRIOR FILING DATE: 1998-07-02  
 PRIOR APPLICATION NUMBER: 60/091626  
 PRIOR FILING DATE: 1998-07-02  
 PRIOR APPLICATION NUMBER: 60/091633  
 PRIOR FILING DATE: 1998-07-02  
 PRIOR APPLICATION NUMBER: 60/091978  
 PRIOR FILING DATE: 1998-07-07  
 PRIOR APPLICATION NUMBER: 60/091982  
 PRIOR FILING DATE: 1998-07-07  
 PRIOR APPLICATION NUMBER: 60/092182  
 PRIOR FILING DATE: 1998-07-09

Query Match 8 2%; Score 91.5; DB 9; Length 919;

Best Local Similarity 22.8%; Pred. No. 0.58; Indels 75; Gaps 8;

Matches 46; Conservative 26; Mismatches 55; Indels 75; Gaps 8;  
 QY 17 LVYSGSMKVAHVAEPKIEIAIKINAMPKMSYGGIYTPAPYSIIPOGSWNS 76  
 DB 310 VLDKSSMG-----GKDLNRNNOAKHFLD-----TYENSSWG 345  
 QY 77 CV-----ACAVNTIKSDLEIFGRITPVG-----DGIRKHEITVINQMP 115  
 DB 346 MYHFDSTATIVNKLIQIKSSDERNTIMAGLPTY-----PLGGTSICSGIKYAFVIGELHS 401  
 QY 116 Q---AAVITLTDGANNLGMNVPVEVYSIYTNRVNCRHVSFA----- 155  
 DB 402 QLDGSEVLLTDEEDNTASSCIDEVK---OSGAIYHFIAGRADEAVIEMSKITGSGSHF 458

QY 156 ---DDAEKAIIDQIVAINSGS 174  
 DB 459 YVSDPAQNNGILDAFGALTSGN 480

RESULT 5  
 US-09-989-293A-379  
 Sequence 379, Application US/09989293A  
 Patent No. US20020177164A1  
 GENERAL INFORMATION:  
 APPLICANT: Ashkenazi, Avi J.  
 APPLICANT: Baker, Kevin P.  
 APPLICANT: Botstein, David  
 APPLICANT: Deanoeyers, Inc  
 APPLICANT: Eaton, Dan L.  
 APPLICANT: Ferrara, Napoleone  
 APPLICANT: Fong, Sherman  
 APPLICANT: Gerber, Hanspeter  
 APPLICANT: Gerritsen, Mary E.  
 APPLICANT: Goddard, Audrey  
 APPLICANT: Grimaldi, J. Christopher  
 APPLICANT: Gurney, Austin L.  
 APPLICANT: Kijavlin, Ivar J.  
 APPLICANT: Napier, Mary A.  
 APPLICANT: Pan, James  
 APPLICANT: Paoni, Nicholas F.  
 APPLICANT: Roy, Margaret Ann  
 APPLICANT: Stewart, Timothy A.  
 APPLICANT: Tumas, Daniel  
 APPLICANT: Watanabe, Colin K.  
 APPLICANT: Williams, P. Mickey  
 APPLICANT: Wood, William I.  
 APPLICANT: Zhang, Zemin  
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 FILE OF INVENTION: Acids Encoding the Same  
 FILE REFERENCE: P2730P1C66  
 CURRENT APPLICATION NUMBER: US/09/989,293A  
 PRIOR FILING DATE: 2001-11-20  
 PRIOR APPLICATION NUMBER: 60/049787  
 PRIOR FILING DATE: 1997-06-16  
 PRIOR APPLICATION NUMBER: 60/062250  
 PRIOR FILING DATE: 1997-10-17  
 PRIOR APPLICATION NUMBER: 60/065186  
 PRIOR FILING DATE: 1997-11-12  
 PRIOR APPLICATION NUMBER: 60/065311  
 PRIOR FILING DATE: 1997-11-13  
 PRIOR APPLICATION NUMBER: 60/066770  
 PRIOR FILING DATE: 1997-11-24  
 PRIOR APPLICATION NUMBER: 60/075945  
 PRIOR FILING DATE: 1998-02-25  
 PRIOR APPLICATION NUMBER: 60/078910  
 PRIOR FILING DATE: 1998-03-20  
 PRIOR APPLICATION NUMBER: 60/083322  
 PRIOR FILING DATE: 1998-04-28  
 PRIOR APPLICATION NUMBER: 60/084600  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/087106  
 PRIOR FILING DATE: 1998-05-28  
 PRIOR APPLICATION NUMBER: 60/087607  
 PRIOR FILING DATE: 1998-06-02  
 PRIOR APPLICATION NUMBER: 60/087609  
 PRIOR FILING DATE: 1998-06-02  
 PRIOR APPLICATION NUMBER: 60/087759  
 PRIOR FILING DATE: 1998-06-02  
 PRIOR APPLICATION NUMBER: 60/087827  
 PRIOR FILING DATE: 1998-06-03  
 PRIOR APPLICATION NUMBER: 60/088021  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088025  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088026  
 PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER:	60/088028
PRIOR FILING DATE:	1998-06-04
PRIOR APPLICATION NUMBER:	60/088029
PRIOR FILING DATE:	1998-06-04
PRIOR APPLICATION NUMBER:	60/088030
PRIOR FILING DATE:	1998-06-04
PRIOR APPLICATION NUMBER:	60/088033
PRIOR FILING DATE:	1998-06-04
PRIOR APPLICATION NUMBER:	60/088326
PRIOR FILING DATE:	1998-06-04
PRIOR APPLICATION NUMBER:	60/088167
PRIOR FILING DATE:	1998-06-05
PRIOR APPLICATION NUMBER:	60/088217
PRIOR FILING DATE:	1998-06-05
PRIOR APPLICATION NUMBER:	60/088655
PRIOR FILING DATE:	1998-06-09
PRIOR APPLICATION NUMBER:	60/088734
PRIOR FILING DATE:	1998-06-10
PRIOR APPLICATION NUMBER:	60/088738
PRIOR FILING DATE:	1998-06-10
PRIOR APPLICATION NUMBER:	60/088810
PRIOR FILING DATE:	1998-06-10
PRIOR APPLICATION NUMBER:	60/088824
PRIOR FILING DATE:	1998-06-10
PRIOR APPLICATION NUMBER:	60/088826
PRIOR FILING DATE:	1998-06-10
PRIOR APPLICATION NUMBER:	60/088858
PRIOR FILING DATE:	1998-06-11
PRIOR APPLICATION NUMBER:	60/088861
PRIOR FILING DATE:	1998-06-11
PRIOR APPLICATION NUMBER:	60/088767
PRIOR FILING DATE:	1998-06-11
PRIOR APPLICATION NUMBER:	60/089105
PRIOR FILING DATE:	1998-06-12
PRIOR APPLICATION NUMBER:	60/089440
PRIOR FILING DATE:	1998-06-16
PRIOR APPLICATION NUMBER:	60/089512
PRIOR FILING DATE:	1998-06-16
PRIOR APPLICATION NUMBER:	60/089514
PRIOR FILING DATE:	1998-06-16
PRIOR APPLICATION NUMBER:	60/089532
PRIOR FILING DATE:	1998-06-17
PRIOR APPLICATION NUMBER:	60/089538
PRIOR FILING DATE:	1998-06-17
PRIOR APPLICATION NUMBER:	60/089588
PRIOR FILING DATE:	1998-06-17
PRIOR APPLICATION NUMBER:	60/089599
PRIOR FILING DATE:	1998-06-17
PRIOR APPLICATION NUMBER:	60/089600
PRIOR FILING DATE:	1998-06-17
PRIOR APPLICATION NUMBER:	60/089653
PRIOR FILING DATE:	1998-06-17
PRIOR APPLICATION NUMBER:	60/089801
PRIOR FILING DATE:	1998-06-18
PRIOR APPLICATION NUMBER:	60/089907
PRIOR FILING DATE:	1998-06-18
PRIOR APPLICATION NUMBER:	60/089908
PRIOR FILING DATE:	1998-06-18
PRIOR APPLICATION NUMBER:	60/089947
PRIOR FILING DATE:	1998-06-19
PRIOR APPLICATION NUMBER:	60/089948
PRIOR FILING DATE:	1998-06-19
PRIOR APPLICATION NUMBER:	60/089952
PRIOR FILING DATE:	1998-06-19
PRIOR APPLICATION NUMBER:	60/090246
PRIOR FILING DATE:	1998-06-22
PRIOR APPLICATION NUMBER:	60/090252

1	PRIOR FILING DATE: 1998-06-22	
2	PRIOR APPLICATION NUMBER: 60/090254	
3	PRIOR FILING DATE: 1998-06-22	
4	PRIOR APPLICATION NUMBER: 60/090349	
5	PRIOR FILING DATE: 1998-06-23	
6	PRIOR APPLICATION NUMBER: 60/090355	
7	PRIOR FILING DATE: 1998-06-23	
8	PRIOR APPLICATION NUMBER: 60/090429	
9	PRIOR FILING DATE: 1998-06-24	
10	PRIOR APPLICATION NUMBER: 60/090431	
11	PRIOR FILING DATE: 1998-06-24	
12	PRIOR APPLICATION NUMBER: 60/090435	
13	PRIOR FILING DATE: 1998-06-24	
14	PRIOR APPLICATION NUMBER: 60/090444	
15	PRIOR FILING DATE: 1998-06-24	
16	PRIOR APPLICATION NUMBER: 60/090445	
17	PRIOR FILING DATE: 1998-06-24	
18	PRIOR APPLICATION NUMBER: 60/090472	
19	PRIOR FILING DATE: 1998-06-24	
20	PRIOR APPLICATION NUMBER: 60/090535	
21	PRIOR FILING DATE: 1998-06-24	
22	PRIOR APPLICATION NUMBER: 60/090540	
23	PRIOR FILING DATE: 1998-06-24	
24	PRIOR APPLICATION NUMBER: 60/090542	
25	PRIOR FILING DATE: 1998-06-24	
26	PRIOR APPLICATION NUMBER: 60/090557	
27	PRIOR FILING DATE: 1998-06-25	
28	PRIOR APPLICATION NUMBER: 60/090678	
29	PRIOR FILING DATE: 1998-06-25	
30	PRIOR APPLICATION NUMBER: 60/090690	
31	PRIOR FILING DATE: 1998-06-25	
32	PRIOR APPLICATION NUMBER: 60/090694	
33	PRIOR FILING DATE: 1998-06-25	
34	PRIOR APPLICATION NUMBER: 60/090655	
35	PRIOR FILING DATE: 1998-06-25	
36	PRIOR APPLICATION NUMBER: 60/090666	
37	PRIOR FILING DATE: 1998-06-25	
38	PRIOR APPLICATION NUMBER: 60/090682	
39	PRIOR FILING DATE: 1998-06-26	
40	PRIOR APPLICATION NUMBER: 60/090683	
41	PRIOR FILING DATE: 1998-06-26	
42	PRIOR APPLICATION NUMBER: 60/091360	
43	PRIOR FILING DATE: 1998-07-01	
44	PRIOR APPLICATION NUMBER: 60/091478	
45	PRIOR FILING DATE: 1998-07-02	
46	PRIOR APPLICATION NUMBER: 60/091544	
47	PRIOR FILING DATE: 1998-07-01	
48	PRIOR APPLICATION NUMBER: 60/091519	
49	PRIOR FILING DATE: 1998-07-02	
50	PRIOR APPLICATION NUMBER: 60/091626	
51	PRIOR FILING DATE: 1998-07-02	
52	PRIOR APPLICATION NUMBER: 60/091633	
53	PRIOR FILING DATE: 1998-07-02	
54	PRIOR APPLICATION NUMBER: 60/091978	
55	PRIOR FILING DATE: 1998-07-07	
56	PRIOR APPLICATION NUMBER: 60/091982	
57	PRIOR FILING DATE: 1998-07-07	
58	PRIOR APPLICATION NUMBER: 60/092182	
59	PRIOR FILING DATE: 1998-07-09	

Query Match	8.2%	Score	91.5	DB	9	Length	919
Best Local Similarity	22.8%	Pred. No.	0.58				
Matches	46	Conservative	26	Mismatches	55	Indels	75
						Gaps	8

[illegible]

[illegible]

RESULT 6  
US-10-06

```

? Sequence 70 Application US/10065547
? Publication No. US20020182638A1
? GENERAL INFORMATION:
? APPLICANT: Eaton, Dan L.
? APPLICANT: Filvaroff, Ellen
? APPLICANT: Gerritsen, Mary E.
? APPLICANT: Goddard, Audrey
? APPLICANT: Godowski, Paul J.
? APPLICANT: Grimaldi, Christopher J.
? APPLICANT: Gurney, Austin L.
? APPLICANT: Macanabe, Colin K.
? APPLICANT: Wood, William I.
? TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
? FILE REFERENCE: P3230R1C1
? CURRENT APPLICATION NUMBER: US/10/063,547
? CURRENT FILING DATE: 2002-05-02
? Prior Application removed - See File Wrapper or Palm
? NUMBER OF SEQ ID NOS: 170
? SEQ ID NO 70
? LENGTH: 919
? TYPE: PRT
? ORGANISM: Homo Sapien
? US-10-063-547-70

```

Query Match	8.2%	Score 91.5	DB 9	Length 919
Best Local Similarity	22.8%	Pred. No. 0.58		
Matches 46	Conservative 26	Mismatches 55	Indels 75	Gaps 8

```

QY 17 IUDYSGSMNMKIVAUREPTELAKAILIKINAAPMKMSYOGGLTFAFVSIIEQGSMS 76
Db 310 YLDDKSGSMG-----GKRLRMQAAHGFLLQ-----TVNGSVMG 343b
QY 77 CV-----AECVAWTKISDLEIFGRLEPVG-----DGKHEVYINMPR 115b
Db 346 WHPFSDTATVYKLLIQIKSSDERMTLMGLPRY-----PLGTSICSGIKAFQYIELHS 401b
QY 116 Q---AAVILLTDGHNHNGMRYEEVKS1YQTNPNVCFFHVSFA-----155b
Db 402 QLDGSEVILLTLDGEBDNTASSCIDVEK---QSGAIVHFALGRADAEVIEWSKITGSGHF 458b
QY 156 ---DDAEKAILDIQVVAANS 174
Db 459 YVSDERQNNGLIDAFGALTSN 480

```

RESULT 7  
US-09-989-735-379

Sequence 379, Application US/09989755  
Publication No. US20020193299A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi J.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Destrovis, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kijavlin, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2730P1C61  
CURRENT APPLICATION NUMBER: US/09/989,735  
CURRENT FILING DATE: 2001-11-19  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09

;; PRIOR APPLICATION NUMBER: 60/088734  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088738  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088742  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088810  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088824  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088826  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088858  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/088861  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/088876  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/089105  
;; PRIOR FILING DATE: 1998-06-12  
;; PRIOR APPLICATION NUMBER: 60/089440  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089512  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089514  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089532  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089538  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089598  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089599  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089600  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089653  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089801  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089907  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089908  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089947  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/089948  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/089952  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/090246  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090252  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090254  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090349  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090355  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090429  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090431  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090435  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090444  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090445  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090472  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090535

;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090540  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090542  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090557  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090676  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090678  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090690  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090694  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090695  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090696  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090862  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/090863  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/091360  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091478  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091544  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091626  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091633  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091978  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/092182  
;; PRIOR FILING DATE: 1998-07-09

Query Match 8.2%; Score 91.5; DB 9; Length 919;  
Best Local Similarity 22.8%; Pred. No. 0.58;  
Matches 46; Conservative 26; Mismatches 55; Indels 75; Gaps 8;

Qy 17 LVYDGSMMKHAIVREKXILAKEALIKYNAAMPKMSYGGTLTFAPYSVIIPGSGMNS 76  
:::||||| : : : : :  
Db 310 VLDRKSGWG-----GKDRLRNKNQAAKHFLQ-----TVENGSGWVG 345  
Qy 77 CV-----AECVNTTIXSDLEIFGRLTGPG-----DGIKMBETVINOMP 115  
|----- : : : : :  
Db 346 MVHFDSTATYVKNLQIKSSDERNTLWAGLPY-----PLGTSISCSGIKYPQYIGELHS 401  
|----- : : : : :  
Qy 116 Q--AAVILTLTGHNLMGNPVEEVKSYQTNPNVCFFHVSFA----- 155  
: : : : : : : : : : :  
Db 402 QLDSEVLLTLTGDSDNTASCTIDEVK--GSGALVHFIALRAADEAVIEMSKITGSGHF 458  
| : : : : : : : : : : :  
Qy 156 ---DDASKALIIDQIVALSNGS 174  
: : : : : : : : : : :  
Db 459 YVSDAONNGLIDAFALTSNG 480

## RESULT 8

US-09-990-444-379  
; Sequence 379, Application US/09990444  
; Publication No. US20020193300A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Deenoyers, Luc  
; APPLICANT: Baton, Dan L.

APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerriksen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kijavina, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Par, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Matanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2730P1C19  
CURRENT FILING DATE: 2001-11-14  
CURRENT APPLICATION NUMBER: US/09/990,444  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-22  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212

PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090431  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090435  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090444  
PRIOR FILING DATE: 1998-06-24



PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090431

PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090435  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090444  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090472  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090535  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090540  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090542  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090557  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090676  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090678  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090690  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090694  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090695  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090696  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090862  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/090863  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/091360  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091478  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091544  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091519  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091626  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091633  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091978  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/091982  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/092182  
PRIOR FILING DATE: 1998-07-09

Query Match 8.2% Score 91.5; DB 9; Length 919;  
Best Local Similarity 22.8%; Pred. No. 0.58;  
Matches 46; Conservative 26; Mismatches 55; Indels 75; Gaps 8;

QY 17 LVYDSSMMKKAIVAREPEILAKELIKINAMPMKSYOGSLYTFAPYVLIPOGSWNS 76  
DB 310 VLDKSGSMG-----GKDRINRMNQAHHFLQ-----TVENGSMWG 345  
QY 77 CV-----AECAVNTIKSDLEIFGRITPVG-----DGIMHETVINQMP 115  
DB 346 WHFSDSTATIVNKLIQIKSDSRNTLMGLPTY-----PLGTSICSGIKYARQVIGELHS 401  
QY 116 Q--AAVILLTGHNHNGMNPVEYKSYIOTNPNCFHVVSFA----- 155  
DB 402 QLDGSEVLLLTGDBNTASSCIDEVK---QSGAIVHFIALGRAADEAVIEMSKITGSHF 458  
QY 156 ---DDAEKKAIIIDQIVALSNGS 174  
DB 459 YVSDAQNNGNLIDAFGALTSN 480



RESULT 10  
US-09-990-436-379  
Sequence 379, Application US/09990436  
Publication No. US20020198148A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi J.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Deenoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gertlisen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kijavlin, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OR INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
TITLE OR INVENTION: Acids Encoding the Same  
FILE REFERENCE: P2730P1C14  
CURRENT APPLICATION NUMBER: US/09/990,436  
CURRENT FILING DATE: 2001-11-14  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030

PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23

PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090431  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090435  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090444  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090472  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090535  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090540  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090542  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090557  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090676  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090678  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090690  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090694  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090695  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090696  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090862  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/090863  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/09160  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091478  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091544  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091519  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091626  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091633  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091978  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/091982  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/092182  
PRIOR FILING DATE: 1998-07-09

Query Match 8.2%; Score 91.5; DB 9; Length 919;

Best Local Similarity 22.8%; Pred. No. 0.58; 55; Indels 75; Gaps 8;

Matches 46; Conservative 26; Mismatches 55; Indels 75; Gaps 8;

QY 17 LVDSGMMKHAIVREPKTELAKKALIKINAMPKKSYOGGLYTFAPYVSVIIPOGSWNS 76  
DB 310 VLDKSSMG-----GKRLANRNOAKKHLQ-----TYBNSWVG 345  
QY 77 CV-----ACAVNTIKSDLEIFGRLPVG---DGIXMHTVINQMP 115  
DB 346 MYHFDSTATIVNKLIQKSSDERNTIWAAGPTY---PLGTSISGSIKVAFOVIGLHS 401  
QY 116 Q--AAVILLTGDHNLGNMPEVKSITYOTNRVVCCHVVSFA----- 155  
DB 402 QLDGSEVLLTDEBDNTASSCIDEVK---OSGAIVHFIALGRADAENVIMSKITGSGSHF 458

QY 156 ---DDEAGKAIIDQIVAINSGS 174  
DB 459 YVSDEAQNNGILDAFGALTSN 480

RESULT 11  
US-09-991-181-379  
Sequence 379, Application US/09991181  
Publication No. US20020197615A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi J.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Deenoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Guiney, Austin L.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zenith  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: P2730P1CS3  
CURRENT APPLICATION NUMBER: US/09/991,181  
PRIOR FILING DATE: 2001-11-16  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252

PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090431  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090435  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090444  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090472  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090535  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090540  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090542  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090557  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090676  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090678  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090690  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090694  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090695  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090696  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090862  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/090863  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/091360  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091478  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091544  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091519  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091626  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091633  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091978  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/091982  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/092182  
PRIOR FILING DATE: 1998-07-09

Query Match 8.2%; Score 91.5; DB 9; Length 919;  
Best Local Similarity 22.8%; Pred. No. 0.58; 55; Indels 75; Gaps 8;

Matches 46; Conservative 26; Mismatches 55; Indels 75; Gaps 8;

Qy 17 LVDSGMMKGVAVREPKIELAKEALIKINAMPKMSYQGLYTFAPYSVIIFQGSWNS 76  
:::|||||  
Db 310 VLDSGSWG-----GKRLRMNQAKHFLQ-----TVNDSWG 345  
Qy 77 CV-----ACAVNTIKSDLEIFGRLLPVG-----DGKHEVYINQMP 115  
Db 346 MHPDSTATIVNKLIQIKSDERTLMAGLPTY----PLGGTISCGIKVAFQVIGELHS 401

OY 116 Q---AAVILLTGDHNNIGMNPVEYKSIYQTNPNVCFHVSPA----- 155  
DB 402 QUDSEVLLTGDGEDTJASCTIDEVK---QSGAIVHFIALGPAABAVIEMSKITGGSHF 458  
OY 156 ---DDAEGKAIIDQIVALNSGS 174  
DB 459 YVSEDAQNNGLIDAFGALITSGN 480

RESULT 12  
US-09-993-687-379  
; Sequence 379, Application US/09993687  
; Publication No. US20020198149n1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Guiney, Austin L.  
; APPLICANT: Kijavlin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tunas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2730PIC11  
; CURRENT APPLICATION NUMBER: US/09/993,687  
; PRIOR FILING DATE: 2002-11-14  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066770  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/075945  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/084600  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/087106  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087609  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087759  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087827  
; PRIOR FILING DATE: 1998-06-03  
; PRIOR APPLICATION NUMBER: 60/088021

; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088025  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088026  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088028  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088029  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088030  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088033  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088326  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088167  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088202  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088212  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088217  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088655  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: 60/088734  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088738  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088742  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088810  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088824  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088826  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088858  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088861  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088876  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/089105  
; PRIOR FILING DATE: 1998-06-12  
; PRIOR APPLICATION NUMBER: 60/089440  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089512  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089514  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089532  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089538  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089598  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089599  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089600  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089653  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089801  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089907  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089908  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089947  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/089948  
; PRIOR FILING DATE: 1998-06-19

PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090431  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090435  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090444  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090472  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090535  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090540  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090542  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090557  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090676  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090678  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090690  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090694  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090695  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090696  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090862  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/090863  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/091360  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091478  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091544  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091519  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091626  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091633  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091978  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/091982  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/092182  
PRIOR FILING DATE: 1998-07-09

Query Match 8.2%; Score 91.5; DB 9; Length 919;

Best Local Similarity 22.8%; Pred. No. 0.58; Matches 46; Conservative 26; Mismatches 55; Indels 75; Gaps 8;

17 LVYDGSMMGHVAVREKXIELAEKAIKTAAMPKMSYGGGLTYFPADYSVITIPQGSWNS 76  
:::|||||:::|

Db 310 VLDKSGSNG-----GKRLNRNQAANKHFLQ-----TWENGSHVG 345  
Qy 77 CV-----AECVANTIKSDLEIFGLTPVG-----DGKHEETVINCMP 115  
Db 346 MVHEDSTATITVINKLIQIKSPERTITMAGLPTY-----PLGTSICSGIKVAFQYIGELHS 401  
Qy 116 Q--AAVILLTDGHNNGMNPVEEVKSIYQTPNVCNVSFA----- 155  
Db 402 QLDSEVLLTLDGEDNTASSCIDEVK---QSGAIVHFTALGRADAEVIMSKITGSGHF 458  
Qy 156 ---DDAEKAIIDQIVALNGSGS 174  
Db 459 YVSDAQNNGLIDAFGALTSGN 480

## RESULT 13

US-09-989-734-379  
Sequence 379, Application US/09989734  
Publication No. US20030003531A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi J.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gottfredsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kijavijn, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas P.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tuma, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OR INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2730PIC64  
CURRENT APPLICATION NUMBER: US/09/989,734  
PRIOR FILING DATE: 2001-11-19  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02

8.28; Score 91.5; DB 9; Length 919;

PRIOR FILING DATE: 1998-05-28	PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02	PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02	PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02	PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02	PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03	PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04	PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04	PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04	PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04	PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04	PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04	PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04	PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04	PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05	PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05	PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05	PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05	PRIOR APPLICATION NUMBER: 60/088555
PRIOR FILING DATE: 1998-06-09	PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10	PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10	PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10	PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10	PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10	PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10	PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11	PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11	PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11	PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12	PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16	PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16	PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16	PRIOR APPLICATION NUMBER: 60/089516
PRIOR FILING DATE: 1998-06-17	PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17	PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17	PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17	PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17	PRIOR APPLICATION NUMBER: 60/089633
PRIOR FILING DATE: 1998-06-17	PRIOR APPLICATION NUMBER: 60/089653

PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090431  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090435  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090444  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090472  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090535  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090540  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090542  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090557  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090676  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090678  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090690  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090694  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090695  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090696  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090862  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/090863  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/091360  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091478  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091544  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091519  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091626  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091633  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091978  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/091982

PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/092182  
PRIOR FILING DATE: 1998-07-09

Query Match  
Best Local Similarity 22.8%; Pred. No. 0.58; 55; Indels 75; Gaps 8;  
Matches 46; Conservative 26; Mismatches 55; Indels 75; Gaps 8;

17 LVDYSGMMKGVAVREPKIELAKEALIKINAMPKMSYGGLYTFAPYSVILPQGSNS 76  
310 VLDSKSGSWG-----GKDRLRNMQAQKHFLLQ-----TWNGSWVG 345

77 CV-----AECAVNTIKSDELIFGRLPVG-----DGIMHETVINQMP 115  
346 WHFEPSTATIVKLIQIKSSDERNTLMAGLPY-----PLGGSISCSGIXKAFQYIGELHS 401

116 Q---AAVILLTDGHNHNGMNPVEEYKSIYQTNPNVCFHVSFA----- 155  
402 QLDGSEVLLTGDENNTASSCIDEVK---QSGAIVHFALGRAADEAVIEMSKITGSHF 458

156 ---DDAEGKAIIDQIVALSNGS 174  
459 YVSDAQNNGLIDARGALTSGN 480

RESULT 15  
US-10-174-590-258  
Sequence 258, Application US/10174590  
Publication No. US2003008352A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jitan  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE OF INVENTION: ACIDS ENCODING THE SAME  
FILE REFERENCE: P3430R1C42  
CURRENT APPLICATION NUMBER: US/10/174,590  
CURRENT FILING DATE: 2002-06-18  
Prior application removed - See file wrapper or Palm  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 258  
LENGTH: 919  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-174-590-258

Query Match  
Best Local Similarity 8.2%; Score 91.5; DB 9; Length 919;  
Best Local Similarity 22.8%; Pred. No. 0.58;  
Matches 46; Conservative 26; Mismatches 55; Indels 75; Gaps 8;

17 LVDYSGMMKGVAVREPKIELAKEALIKINAMPKMSYGGLYTFAPYSVILPQGSNS 76  
310 VLDSKSGSWG-----GKDRLRNMQAQKHFLLQ-----TWNGSWVG 345

77 CV-----AECAVNTIKSDELIFGRLPVG-----DGIMHETVINQMP 115  
346 WHFEPSTATIVKLIQIKSSDERNTLMAGLPY-----PLGGSISCSGIXKAFQYIGELHS 401

116 Q---AAVILLTDGHNHNGMNPVEEYKSIYQTNPNVCFHVSFA----- 155  
402 QLDGSEVLLTGDENNTASSCIDEVK---QSGAIVHFALGRAADEAVIEMSKITGSHF 458

156 ---DDAEGKAIIDQIVALSNGS 174  
459 YVSDAQNNGLIDARGALTSGN 480



Search completed: April 9, 2003, 14:41:23  
Job time : 40 secs

---



GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 9, 2003, 14:18:16 ; Search time 74 Seconds  
(without alignments)  
392.549 Million cell updates/sec

Title: US-10-034-500-2

Perfect score: 1119  
Sequence: 1 AEVTASCTKRVESYNTLVLDY.....OEQILVTEVVVLRGVNFAR 2.8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1980.DAT.\*  
2: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1981.DAT.\*  
3: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1982.DAT.\*  
4: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1983.DAT.\*  
5: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1984.DAT.\*  
6: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1985.DAT.\*  
7: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1986.DAT.\*  
8: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1987.DAT.\*  
9: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1988.DAT.\*  
10: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1989.DAT.\*  
11: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1990.DAT.\*  
12: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1991.DAT.\*  
13: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1992.DAT.\*  
14: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1993.DAT.\*  
15: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1994.DAT.\*  
16: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1995.DAT.\*  
17: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1996.DAT.\*  
18: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1997.DAT.\*  
19: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1998.DAT.\*  
20: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1999.DAT.\*  
21: /SID2/gcgdata/geneeq/geneeqp-emb1/AA2000.DAT.\*  
22: /SID2/gcgdata/geneeq/geneeqp-emb1/AA2001.DAT.\*  
23: /SID2/gcgdata/geneeq/geneeqp-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	92	8.2	868	AAU29749	Novel human secret
2	92	8.2	906	AAU29751	Novel human secret
3	91.5	8.2	917	AAE22687	Human CS193 Consen
4	91.5	8.2	917	AAU88029	Human calcium-acti
5	91.5	8.2	919	AAV66749	Membrane-bound pro
6	91.5	8.2	919	AAU29152	Human PRO polypept
7	91.5	8.2	919	AAAB7560	Human PRO1124. HC
8	91.5	8.2	931	AAAB5272	Human PRO1124. UNO
9	91.5	8.2	931	ABBI1973	Human membrane-bou
10	90.5	8.1	375	AAU52280	Protonibacterium

11	90	8.0	1091	23	AB848258
12	87.5	7.8	932	23	AB809706
13	87.5	7.8	933	23	AB809707
14	86	7.7	1222	21	AB801830
15	86	7.7	1228	21	AB801838
16	86	7.7	4999	23	AA022158
17	85.5	7.6	1239	21	AAV70062
18	85.5	7.6	1551	22	AA441081
19	85.5	7.6	1588	22	AA439295
20	85.5	7.6	1588	22	AA439295
21	85.5	7.6	1624	23	AAE33979
22	84	7.5	488	21	AA818447
23	84	7.5	488	21	AA818455
24	83.5	7.5	509	19	AAW77323
25	83	7.4	351	21	AA621301
26	83	7.4	362	21	AA621300
27	83	7.4	488	21	AA818456
28	83	7.4	488	21	AA818457
29	83	7.4	579	21	AA621299
30	82.5	7.4	689	21	AA647452
31	82.5	7.4	700	21	AA647451
32	82.5	7.4	917	21	AA647450
33	82	7.3	913	20	AA874832
34	82	7.3	913	22	AA873715
35	82	7.3	913	23	AA888027
36	81	7.2	886	23	AA857133
37	81	7.2	2183	19	AAW48707
38	81	7.2	2183	19	AAW48708
39	81	7.2	2183	19	AAW48709
40	81	7.2	2183	19	AAW48703
41	81	7.2	2183	19	AAW48704
42	80.5	7.2	786	22	AA882047
43	80.5	7.2	786	23	AA015400
44	80.5	7.2	957	23	AAE33375
45	80.5	7.2	1397	22	ABG15812

#### ALIGNMENTS

RESULT 1  
AAU29749  
ID AAU29749 standard; Protein; 868 AA.  
XX  
AC AAU29749;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Novel human secreted protein #240.  
XX  
KW Human; vaccination; gene therapy; nutritional supplement;  
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
XX  
OS Homo sapiens.  
XX  
PN WO200179449-A2.  
XX  
PD 25-OCT-2001.  
XX  
PF 16-APR-2001; 2001WO-US08656.  
XX  
PR 18-APR-2000; 2000US-0552929.  
PR 26-JAN-2001; 2001US-0770160.  
XX  
PA (HYSE-) HYSEO INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
DR WPI; 2001-611725/70.  
XX  
PT Nucleic acids encoding a range of human polypeptides, useful in genetic  
vaccination, testing and therapy -

XX Claim 20; Page 189; 765pp; English.  
XX  
XX The invention relates to novel human secreted polypeptides. The  
CC polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC expressing the proteins are useful for identifying a therapeutic agent  
CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising  
CC the nucleic acids encoding the polypeptides and cells genetically  
CC engineered to express them are also useful for producing the proteins.  
CC The proteins are useful in genetic vaccination, testing and  
CC therapy, and can be used as nutritional supplements. They may be used to  
CC increase stem cell proliferation; to regulate haematopoiesis; and in  
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;  
CC immune suppression and/or stimulation; as anti-inflammatory agents; and  
CC in treatment of leukaemias. AAU29510-AAU3304 represent the amino acid  
CC sequences of novel human secreted proteins of the invention.  
XX  
SQ Sequence 868 AA;  
Query Match 8.2%; Score 92; DB 22; Length 868;  
Best Local Similarity 24.8%; Pred. No. 0.89;  
Matches 34; Conservative 30; Mismatches 45; Indels 28; Gaps 7;  
QY 16 YLVVDYSGSMKMHVAVREPKIEFLAEALIKINAAMPKMSYOGGLTFAPYSVILPQGSWN 75  
DB 280 FVIDISGSMAGR-----KLEQTEALRLILEDKEDY---LNFLFSGDV--STWK 326  
QY 76 SCVAECA-----VNTIKSDEIFGRLLTPVGDGIMKHEVTIN-----QMPPOAA--VIL 121  
DB 327 EHLVQATPENIQEARTFVKSMEDKG-MTNINDGLLRGISMLKAKBEHRIPERSYSTIVIM 385  
QY 122 LTGHNNTIGANPVEEVK 138  
DB 386 LTGDDANVGESRPEKIQ 402  
RESULT 2  
AAU29751  
ID AAU29751 standard; Protein; 906 AA.  
XX  
XX AAU29751;  
AC  
XX 18-DEC-2001 (first entry)  
DT  
XX  
XX Novel human secreted protein #242.  
DE  
XX  
XX Human; vaccination; gene therapy; nutritional supplement;  
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200179449-A2.  
PN  
XX  
XX 25-OCT-2001.  
PD  
XX  
XX 16-APR-2001; 2001WO-US08656.  
PF  
XX  
XX 18-APR-2000; 2000US-0552929.  
PR  
XX 26-JAN-2001; 2001US-0770160.  
PR  
XX  
XX (HYSE-) HYSEQ INC.  
PA  
XX  
XX Tang YT, Liu C, Drmanac RT;  
PI  
XX  
XX WPI; 2001-611725/70.  
DR  
XX  
XX Nucleic acids encoding a range of human polypeptides, useful in genetic  
PT vaccination, testing and therapy -

XX Claim 20; Page 189; 765pp; English.  
XX  
XX The invention relates to novel human secreted polypeptides. The  
CC polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC expressing the proteins are useful for identifying a therapeutic agent  
CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising  
CC the nucleic acids encoding the polypeptides and cells genetically  
CC engineered to express them are also useful for producing the proteins.  
CC The proteins are useful in genetic vaccination, testing and  
CC therapy, and can be used as nutritional supplements. They may be used to  
CC increase stem cell proliferation; to regulate haematopoiesis; and in  
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;  
CC immune suppression and/or stimulation; as anti-inflammatory agents; and  
CC in treatment of leukaemias. AAU29510-AAU3304 represent the amino acid  
CC sequences of novel human secreted proteins of the invention.  
XX  
SQ Sequence 906 AA;  
Query Match 8.2%; Score 92; DB 22; Length 906;  
Best Local Similarity 24.8%; Pred. No. 0.94;  
Matches 34; Conservative 30; Mismatches 45; Indels 28; Gaps 7;  
QY 16 YLVVDYSGSMKMHVAVREPKIEFLAEALIKINAAMPKMSYOGGLTFAPYSVILPQGSWN 75  
DB 289 FVIDISGSMAGR-----KLEQTEALRLILEDKEDY---LNFLFSGDV--STWK 335  
QY 76 SCVAECA-----VNTIKSDEIFGRLLTPVGDGIMKHEVTIN-----QMPPOAA--VIL 121  
DB 336 EHLVQATPENIQEARTFVKSMEDKG-MTNINDGLLRGISMLKAKBEHRIPERSYSTIVIM 394  
QY 122 LTGHNNTIGANPVEEVK 138  
DB 395 LTGDDANVGESRPEKIQ 411  
RESULT 3  
AAE22687  
ID AAE22687 standard; Protein; 917 AA.  
XX  
XX AAE22687;  
AC  
XX  
XX 26-JUL-2002 (first entry)  
DT  
XX  
XX Human CS193 consensus sequence polypeptide.  
DE  
XX  
XX Human; CS193; gastrointestinal tract; cancer; gene therapy;  
KW cytostatic.  
KW  
XX  
XX Homo sapiens.  
OS  
XX  
XX US6368792-B1.  
PN  
XX  
XX 09-APR-2002.  
PD  
XX  
XX 27-MAR-1998; 98US-0049698.  
PF  
XX  
XX 31-MAR-1997; 97US-0828856.  
PR  
XX  
XX (ABBO ) ABBOTT LAB.  
PA  
XX  
XX Billiguel PA, Cohen M, Colpitta TL, Friedman PM, Hayden M;  
PI Klaas MR, Roberts-rapp L, Russell JC, Stroupe SD;  
XX  
XX WPI; 2002-328082/36.  
DR  
XX N-PSDB; AAD35925.  
DR  
XX  
XX New purified polynucleotide encoding CS193 antigen, useful for  
PT diagnosing, staging, monitoring preventing or treating gastrointestinal

PT disorders -  
XX  
XX Example 1; Column 87-92; 58pp; English.  
PS  
XX The invention relates to a purified polynucleotide encoding CS193. The  
CC polynucleotide is used for detecting, diagnosing, staging, monitoring,  
CC prognosticating, preventing or treating diseases and conditions of the  
CC gastrointestinal tract, particularly cancer. The CS193 gene is useful in  
CC gene therapy. The present sequence is human CS193 consensus sequence  
CC polypeptide.  
XX  
XX Sequence 917 AA;  
Query Match 8.2%; Score 91.5; DB 23; Length 917;  
Best Local Similarity 22.8%; Pred. No. 1.1;  
Matches 46; Conservative 26; Mismatches 55; Indels 75; Gaps 8;  
QY 17 LVDSGSMGMHVAVREKIELAEALIKINAAAPKMSYGGGLYFAPYVYIIPQGSNNS 76  
DB 310 VLDSGSMG-----GKDRLRNMQAKHFLQ-----TWENGSMVG 345  
QY 77 CV-----AECAVNTIKSDLEIFGRTPVG-----DGIKMHEVINOMP 115  
DB 346 MVHFDSTATYVVKLIQIKSSDERNTIMAGLPY-----PLGGISISGIKYAFQVIGELHS 401  
QY 116 Q---AAVILLTDGHNHNLGAMPVEEYKSIYQTNPNVCFHVVSFA----- 155  
DB 402 QLDGSEVLLTLDGEDNTASSCIDEYK---QSGAIVHFYALGRADEAVIEMSKITGSGHF 458  
QY 156 ---DDAEGKAIIDQIVALNSGS 174  
DB 459 YVSDAQNNGLIDAFGALTSGN 480  
RESULT 4  
ID AAU88029 standard; Protein; 917 AA.  
XX  
AC AAU88029;  
XX  
DT 05-JUN-2002 (first entry)  
XX  
DE Human calcium-activated chloride channel hCLCA4.  
XX  
XX Nucleic acid library; immune response; asthma; COPD;  
KM airway hyperresponsiveness; bronchoalveolar manifestation;  
KM signature sequence; SS: chronic obstructive pulmonary disease;  
KM allergic disease; rhinitis; atopic dermatitis; urticaria;  
KM autoimmune disease; multiple sclerosis; inflammatory bowel disease;  
KM allograft rejection; infectious disease.  
KM calcium-activated chloride channel.  
OS  
XX Homo sapiens.  
XX  
XX WO200214366-A2.  
XX  
XX 21-FEB-2002.  
XX  
XX 16-AUG-2001; 2001WO-NL00610.  
XX  
XX 16-AUG-2000; 2000EP-0202867.  
XX  
XX (UYUT-) RIKJSUNIV UTRCHT.  
XX  
XX Grooc PC, Van Bergenhenegouwen BJ, Van Oosterhout AJM;  
XX  
XX WPI; 2002-241888/29.  
XX  
XX Nucleic acid library comprising genes which are capable of initiation,  
PT progression and suppression of an immune response, especially an immune  
PT response observed with airway hyper-responsiveness of asthma -  
XX  
XX Disclosure; Fig 14; 120pp; English.  
PS

XX  
CC The invention relates to a nucleic acid library comprising genes or  
CC their fragments which are capable of modulating an immune response  
CC observed with airway hyperresponsiveness and/or bronchoalveolar  
CC manifestations of asthma. Also included are a method for modulating an  
CC immune response of an individual comprising modulating a gene comprising  
CC a nucleic acid at least functionally equivalent to a nucleic acid  
CC identifiable by a signature sequence (SS) given in the specification such  
CC as R1-SO-R1-A11, SC01-A10, SVO2-1-C11, SC01-A12, and R1-SO-R1-B7, a  
CC substance (for use as a medicament) capable of modulating a gene  
CC comprising a nucleic acid at least functionally equivalent to a nucleic  
CC acid identifiable by SS and the use of a proteaceous substance derived  
CC from a nucleic acid at least functionally equivalent to a nucleic acid  
CC identifiable by SS for the production of an antagonist (for use as a  
CC medicament) against the substance. The antagonist and substance are  
CC useful for the treatment of an immune response observed with airway  
CC hyperresponsiveness and/or bronchoalveolar manifestations of asthma.  
CC The method is useful for modulating the above immune response, where the  
CC gene encodes a gene product capable of modulating the immune response.  
CC The substance is useful for treating an immune response, particularly  
CC asthma, chronic obstructive pulmonary disease (COPD), allergic diseases  
CC (rhinitis, atopic dermatitis, urticaria), autoimmune diseases (e.g.  
CC multiple sclerosis), inflammatory bowel disease, allograft rejection and  
CC infectious disease. The present sequence is a mouse or human  
CC protein encoded by a signature sequence gene or its homologue/functional  
CC equivalent.  
XX  
XX Sequence 917 AA;  
Query Match 8.2%; Score 91.5; DB 23; Length 917;  
Best Local Similarity 22.8%; Pred. No. 1.1;  
Matches 46; Conservative 26; Mismatches 55; Indels 75; Gaps 8;  
QY 17 LVDSGSMGMHVAVREKIELAEALIKINAAAPKMSYGGGLYFAPYVYIIPQGSNNS 76  
DB 310 VLDSGSMG-----GKDRLRNMQAKHFLQ-----TWENGSMVG 345  
QY 77 CV-----AECAVNTIKSDLEIFGRTPVG-----DGIKMHEVINOMP 115  
DB 346 MVHFDSTATYVVKLIQIKSSDERNTIMAGLPY-----PLGGISISGIKYAFQVIGELHS 401  
QY 116 Q---AAVILLTDGHNHNLGAMPVEEYKSIYQTNPNVCFHVVSFA----- 155  
DB 402 QLDGSEVLLTLDGEDNTASSCIDEYK---QSGAIVHFYALGRADEAVIEMSKITGSGHF 458  
QY 156 ---DDAEGKAIIDQIVALNSGS 174  
DB 459 YVSDAQNNGLIDAFGALTSGN 480  
RESULT 5  
ID AAY66749 standard; Protein; 919 AA.  
XX  
AC AAY66749;  
XX  
XX 05-APR-2000 (first entry)  
XX  
XX Membrane-bound protein PRO1124.  
XX  
XX Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIR ligand;  
KM pharmaceutical; receptor immunoadhesin; gene mapping.  
XX  
XX Homo sapiens.  
XX  
XX WO9963088-A2.  
XX  
XX 09-DEC-1999.  
XX  
XX 02-JUN-1999; 99WO-US12252.  
XX  
XX 02-JUN-1998; 98US-0087607.  
XX  
XX 02-JUN-1998; 98US-0087609.  
PR

PR 02-JUN-1998; 98US-0087759.  
PR 03-JUN-1998; 98US-0087827.  
PR 04-JUN-1998; 98US-0088021.  
PR 04-JUN-1998; 98US-0088025.  
PR 04-JUN-1998; 98US-0088028.  
PR 04-JUN-1998; 98US-0088029.  
PR 04-JUN-1998; 98US-0088030.  
PR 04-JUN-1998; 98US-0088033.  
PR 04-JUN-1998; 98US-0088126.  
PR 05-JUN-1998; 98US-0088167.  
PR 05-JUN-1998; 98US-0088202.  
PR 05-JUN-1998; 98US-0088212.  
PR 05-JUN-1998; 98US-0088217.  
PR 09-JUN-1998; 98US-0088655.  
PR 10-JUN-1998; 98US-0088722.  
PR 10-JUN-1998; 98US-0088730.  
PR 10-JUN-1998; 98US-0088734.  
PR 10-JUN-1998; 98US-0088738.  
PR 10-JUN-1998; 98US-0088740.  
PR 10-JUN-1998; 98US-0088741.  
PR 10-JUN-1998; 98US-0088742.  
PR 10-JUN-1998; 98US-0088810.  
PR 10-JUN-1998; 98US-0088811.  
PR 10-JUN-1998; 98US-0088824.  
PR 10-JUN-1998; 98US-0088825.  
PR 10-JUN-1998; 98US-0088826.  
PR 10-JUN-1998; 98US-0088858.  
PR 11-JUN-1998; 98US-0088861.  
PR 11-JUN-1998; 98US-0088863.  
PR 11-JUN-1998; 98US-0088876.  
PR 12-JUN-1998; 98US-0089090.  
PR 12-JUN-1998; 98US-0089105.  
PR 16-JUN-1998; 98US-0089440.  
PR 16-JUN-1998; 98US-0089512.  
PR 16-JUN-1998; 98US-0089514.  
PR 17-JUN-1998; 98US-0089532.  
PR 17-JUN-1998; 98US-0089538.  
PR 17-JUN-1998; 98US-0089598.  
PR 17-JUN-1998; 98US-0089599.  
PR 17-JUN-1998; 98US-0089600.  
PR 17-JUN-1998; 98US-0089653.  
PR 18-JUN-1998; 98US-0089801.  
PR 18-JUN-1998; 98US-0089907.  
PR 18-JUN-1998; 98US-0089908.  
PR 19-JUN-1998; 98US-0089947.  
PR 19-JUN-1998; 98US-0089948.  
PR 19-JUN-1998; 98US-0089952.  
PR 22-JUN-1998; 98US-0090246.  
PR 22-JUN-1998; 98US-0090252.  
PR 22-JUN-1998; 98US-0090254.  
PR 23-JUN-1998; 98US-0090349.  
PR 23-JUN-1998; 98US-0090355.  
PR 24-JUN-1998; 98US-0090429.  
PR 24-JUN-1998; 98US-0090431.  
PR 24-JUN-1998; 98US-0090435.  
PR 24-JUN-1998; 98US-0090444.  
PR 24-JUN-1998; 98US-0090445.  
PR 24-JUN-1998; 98US-0090461.  
PR 24-JUN-1998; 98US-0090472.  
PR 24-JUN-1998; 98US-0090535.  
PR 24-JUN-1998; 98US-0090538.  
PR 24-JUN-1998; 98US-0090540.  
PR 24-JUN-1998; 98US-0090557.  
PR 25-JUN-1998; 98US-0090676.  
PR 25-JUN-1998; 98US-0090678.  
PR 25-JUN-1998; 98US-0090688.  
PR 25-JUN-1998; 98US-0090690.  
PR 25-JUN-1998; 98US-0090691.  
PR 25-JUN-1998; 98US-0090694.  
PR 25-JUN-1998; 98US-0090695.  
PR 25-JUN-1998; 98US-0090696.  
PR 26-JUN-1998; 98US-0090862.  
PR 26-JUN-1998; 98US-0090863.  
  
PR 01-JUL-1998; 98US-0091358.  
PR 01-JUL-1998; 98US-0091360.  
PR 01-JUL-1998; 98US-0091544.  
PR 02-JUL-1998; 98US-0091478.  
PR 02-JUL-1998; 98US-0091486.  
PR 02-JUL-1998; 98US-0091519.  
PR 02-JUL-1998; 98US-0091628.  
PR 02-JUL-1998; 98US-0091628.  
PR 02-JUL-1998; 98US-0091633.  
PR 02-JUL-1998; 98US-0091646.  
PR 02-JUL-1998; 98US-0091673.  
PR 07-JUL-1998; 98US-0091978.  
PR 07-JUL-1998; 98US-0091982.  
PR 09-JUL-1998; 98US-0092182.  
PR 10-JUL-1998; 98US-0092472.  
PR 20-JUL-1998; 98US-0093339.  
PR 30-JUL-1998; 98US-0094651.  
PR 04-AUG-1998; 98US-0095282.  
PR 04-AUG-1998; 98US-0095285.  
PR 04-AUG-1998; 98US-0095301.  
PR 04-AUG-1998; 98US-0095302.  
PR 04-AUG-1998; 98US-0095318.  
PR 04-AUG-1998; 98US-0095321.  
PR 04-AUG-1998; 98US-0095325.  
PR 10-AUG-1998; 98US-0095916.  
PR 10-AUG-1998; 98US-0095929.  
PR 10-AUG-1998; 98US-0096012.  
PR 11-AUG-1998; 98US-0096143.  
PR 11-AUG-1998; 98US-0096146.  
PR 12-AUG-1998; 98US-0096339.  
PR 17-AUG-1998; 98US-0096757.  
PR 17-AUG-1998; 98US-0096766.  
PR 17-AUG-1998; 98US-0096773.  
PR 17-AUG-1998; 98US-0096791.  
PR 17-AUG-1998; 98US-0096867.  
PR 17-AUG-1998; 98US-0096891.  
PR 17-AUG-1998; 98US-0096894.  
PR 17-AUG-1998; 98US-0096895.  
PR 17-AUG-1998; 98US-0096897.  
PR 18-AUG-1998; 98US-0096949.  
PR 18-AUG-1998; 98US-0096950.  
PR 18-AUG-1998; 98US-0096953.  
PR 18-AUG-1998; 98US-0096960.  
PR 18-AUG-1998; 98US-0097022.  
PR 19-AUG-1998; 98US-0097141.  
PR 20-AUG-1998; 98US-0097218.  
PR 24-AUG-1998; 98US-0097661.  
PR 26-AUG-1998; 98US-0097951.  
PR 26-AUG-1998; 98US-0097952.  
PR 26-AUG-1998; 98US-0097954.  
PR 26-AUG-1998; 98US-0097955.  
PR 26-AUG-1998; 98US-0097971.  
PR 26-AUG-1998; 98US-0097974.  
PR 26-AUG-1998; 98US-0097978.  
PR 26-AUG-1998; 98US-0097979.  
PR 26-AUG-1998; 98US-0097986.  
PR 26-AUG-1998; 98US-0098014.  
PR 31-AUG-1998; 98US-0098525.  
PR 16-SEP-1998; 98US-0100634.  
PR 12-JAN-1999; 99US-0115563.  
  
PA (GETH ) GENENTECH INC.  
XX  
XX Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK;  
PI Wood WI, Yuan J;  
XX  
XX MPI: 2000-072883/06.  
DR N-PSDB; AAZ65095.  
XX  
XX Membrane-bound proteins and related nucleotide sequences -  
PS claim 12; Fig 274; 822pp; English.

XX The invention provides membrane-bound PRO polypeptides and  
 CC polynucleotides encoding them. The PRO sequences of the invention were  
 CC identified based on extracellular domain homology screening. The PRO  
 CC sequences have homology with proteins including LDL receptors, TIE  
 CC ligands and various enzymes. The membrane-bound proteins and receptor  
 CC molecules are useful as pharmaceutical and diagnostic agents. Receptor  
 CC immunoadhesins, for instance, can be used as therapeutic agents to block  
 CC receptor-ligand interactions. The membrane-bound proteins can also be  
 CC employed for screening of potential peptide or small molecule inhibitors  
 CC of the relevant receptor/ligand interaction. The PRO encoding sequences  
 CC are useful as hybridization probes, in chromosome and gene mapping and in  
 CC the generation of antisense RNA and DNA. PRO nucleic acid sequences  
 CC will also be useful for the preparation of PRO polypeptides, especially  
 CC by recombinant techniques.

XX Sequence 919 AA;  
 SQ

Query Match 8.2%; Score 91.5; DB 21; Length 919;  
 Best Local Similarity 22.8%; Pred. No. 1.1;  
 Matches 46; Conservative 26; Mismatches 55; Indels 75; Gaps 8;

QY 17 LVVDSGSMGMHVAVREKIELAKEALIKINAMPKMSYGGGLTFAPYSVILPQSGWNS 76  
 Db 310 VLKSGSGWG-----GKDRLRNMQAKHFLDQ-----TWENGSVWG 345  
 QY 77 CV-----AECVAVTIKSDLEIFGRLETPVG-----DGIKMHEVYINQMP 115  
 Db 346 MVHFDSTATTIVNKLIQIKSDERTITLMAGLPTY-----PLGGTSTICSGIKVAFVYIGELHS 401  
 QY 116 Q--AAVILLTDGHNINIGMNPVEEVKSIYQTNPNVCFFHVSFA----- 155  
 Db 402 QLDSEVILLTDGHDNFTASSCIDEVK---QSGAIVHFIALGRAADEAVIEMSKITGSGHF 458  
 QY 156 ---DDAEKALIIDQIVALNSGS 174  
 Db 459 YVSDAQNNGLIDAFALTSN 480

RESULT 6  
 AAU29152  
 ID AAU29152 standard; Protein; 919 AA.  
 AC AAU29152;  
 XX  
 DT 18-DEC-2001 (first entry)  
 XX  
 DE Human PRO polypeptide sequence #129.  
 XX  
 KM PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;  
 KM dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;  
 KM blood; chondrocyte cell; cell proliferation; cell differentiation; colon;  
 KM adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO200168848-A2.  
 XX  
 PD 20-SEP-2001.  
 XX  
 PF 28-FEB-2001; 2001WO-US06520.  
 XX  
 PR 01-MAR-2000; 2000WO-US05601.  
 PR 02-MAR-2000; 2000WO-US05841.  
 PR 03-MAR-2000; 2000US-187202P.  
 PR 06-MAR-2000; 2000US-186968P.  
 PR 14-MAR-2000; 2000US-189328P.  
 PR 15-MAR-2000; 2000US-189328P.  
 PR 15-MAR-2000; 2000WO-US06884.  
 PR 21-MAR-2000; 2000US-190828P.  
 PR 21-MAR-2000; 2000US-191007P.  
 PR 21-MAR-2000; 2000US-191048P.  
 PR 21-MAR-2000; 2000US-191314P.

PR 28-MAR-2000; 2000US-192655P.  
 PR 29-MAR-2000; 2000US-193032P.  
 PR 29-MAR-2000; 2000US-193033P.  
 PR 30-MAR-2000; 2000WO-US08439.  
 PR 04-APR-2000; 2000US-194443P.  
 PR 04-APR-2000; 2000US-194647P.  
 PR 11-APR-2000; 2000US-195975P.  
 PR 11-APR-2000; 2000US-196000P.  
 PR 11-APR-2000; 2000US-196187P.  
 PR 11-APR-2000; 2000US-196690P.  
 PR 11-APR-2000; 2000US-196820P.  
 PR 18-APR-2000; 2000US-198121P.  
 PR 18-APR-2000; 2000US-198585P.  
 PR 25-APR-2000; 2000US-199397P.  
 PR 25-APR-2000; 2000US-199550P.  
 PR 25-APR-2000; 2000US-199654P.  
 PR 03-MAY-2000; 2000US-201515P.  
 PR 17-MAY-2000; 2000WO-US13705.  
 PR 22-MAY-2000; 2000WO-US14042.  
 PR 30-MAY-2000; 2000WO-US14941.  
 PR 02-JUN-2000; 2000WO-US15264.  
 PR 05-JUN-2000; 2000US-209832P.  
 PR 28-JUL-2000; 2000WO-US20710.  
 PR 22-AUG-2000; 2000US-064848.  
 PR 24-AUG-2000; 2000WO-US23328.  
 PR 08-NOV-2000; 2000WO-US30952.  
 PR 01-DEC-2000; 2000WO-US32678.  
 PR 20-DEC-2000; 2000WO-US34956.

(GETH ) GENENTECH INC.  
 XX  
 XX Baker KP, Chen J, Deenoyers L, Goddard A, Godowski PJ, Gurney AL;  
 PI Pan J, Smith V, Watanabe CK, Wood WJ, Zhang Z;  
 DR WPI; 2001-602746/68.  
 DR N-PSDB; AAS46053.  
 XX  
 PT Novel nucleic acids encoding PRO polypeptides, used to diagnose the  
 PT presence of tumours, such as prostate and breast tumours, in mammals and  
 PT to screen for modulators of the compounds -  
 XX  
 PS Claim 11; Fig 258; 774pp; English.  
 XX  
 CC Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.  
 CC The PRO polypeptides and their associated nucleic acids can be used to  
 CC detect the presence of a tumour in a mammal by comparing the level of  
 CC expression of a PRO polypeptide in a test sample of cells from the animal  
 CC and a control sample of normal cells, whereby a higher level of  
 CC expression in the test sample indicates the presence of a tumour in the  
 CC mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats  
 CC and rabbits but are preferably human. The polypeptides can be used to  
 CC stimulate tumour necrosis factor (TNF) alpha release from human blood,  
 CC when contacted with it. A specific polypeptide can be used to stimulate  
 CC the proliferation or differentiation of chondrocyte cells. The PRO  
 CC proteins can be used to determine the presence of tumours and also  
 CC susceptibility to tumour development, particularly adrenal, lung, colon,  
 CC breast, prostate, rectal, cervical, or liver tumours, in mammalian  
 CC subjects. The oligonucleotide probes specific for the PRO nucleic acids  
 CC can be used for genetic analysis of individuals with genetic disorders.  
 XX

SQ Sequence 919 AA;  
 SQ

Query Match 8.2%; Score 91.5; DB 22; Length 919;  
 Best Local Similarity 22.8%; Pred. No. 1.1;  
 Matches 46; Conservative 26; Mismatches 55; Indels 75; Gaps 8;

QY 17 LVVDSGSMGMHVAVREKIELAKEALIKINAMPKMSYGGGLTFAPYSVILPQSGWNS 76  
 Db 310 VLKSGSGWG-----GKDRLRNMQAKHFLDQ-----TWENGSVWG 345  
 QY 77 CV-----AECVAVTIKSDLEIFGRLETPVG-----DGIKMHEVYINQMP 115  
 Db 346 MVHFDSTATTIVNKLIQIKSDERTITLMAGLPTY-----PLGGTSTICSGIKVAFVYIGELHS 401



```

Oy 116 Q---AAVILLTDGHNLMGNPVEVKSIYQTPNVCFFHVSFA----- 155
Db 402 QLDGSEVLLITDGBDNTSSCIDEVK---QSGALVHFTALGRAADENAVIEMSKITGSGHF 458
Oy 156 ---DDAEGKAIIDQIVAINSGS 174
Db 459 YVSDEAQNNGILDAFGALTSGN 480

RESULT 7
AAB87560
ID AAB87560 standard; Protein; 919 AA.
XX
AC AAB87560;
XX
DT 15-MAY-2001 (first entry)
XX
DE Human PRO1124.
XX
KW Human; PRO protein; mapping.
XX
OS Homo sapiens.
XX
PN MO200116318-A2.
XX
PD 08-MAR-2001.
XX
PF 24-AUG-2000; 2000WO-US23328.
XX
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 07-DEC-1999; 99US-0169495.
PR 09-DEC-1999; 99US-0170262.
PR 11-JAN-2000; 2000US-0175481.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 01-MAR-2000; 2000WO-US05601.
PR 03-MAR-2000; 2000US-0187202.
PR 25-APR-2000; 2000US-0199397.
PR 22-MAY-2000; 2000WO-US14042.
PR 05-JUN-2000; 2000US-0209832.
XX
PA (GETH ) GENENTECH INC.
XX
PI Eaton DL, Filvarov E, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi CJ, Gurney AL, Watanabe CK, Wood WI;
XX
DR MPI; 2001-183260/18.
DR N-PSDB; AAF92092.
XX
PT Eighty four nucleic acids encoding PRO polypeptides, useful in
PT molecular biology, including use as hybridization probes, and in
PT chromosome and gene mapping.
XX
PS Claim 12; Fig 70; 278pp; English.
XX
CC The present sequence is a human PRO polypeptide (secreted and
CC transmembrane). The PRO protein, and PRO agonists, PRO antagonists or
CC anti-PRO antibodies are useful for preparation of a medicament useful in
CC the treatment of a condition which is responsive to the PRO protein,
CC agonists, antagonists or anti-PRO antibodies. The PRO protein may also be
CC employed as molecular weight markers for protein electrophoresis. The PRO
CC coding sequence has applications in molecular biology, including use as
CC hybridisation probes, and in chromosome and gene mapping.
XX
SQ Sequence 919 AA;
Query Match 8.2%; Score 91.5; DB 22; Length 919;
Best Local Similarity 22.8%; Pred. No. 1.1;
Matches 46; Conservative 26; Mismatches 55; Indels 75; Gaps 8;

```

```

Oy 17 LVDSGSMNMKHAVAREPKLEALAEALIKINAMPKMSYGGGLTFAFYSIITDGSNS 76
Db 310 VLDSGSGSWG-----GKRLNRMNQAKHFLIQ-----TVENGSGWG 345
Oy 77 CV-----AECANVTIKSDLEIFGRILPVG-----DGIKHETVYNQMP 115
Db 346 MHPDSTATIVNKLIIQIKSDERTLMAGLPY---PLGGTSCSGIKVAFVYIGELHS 401
Oy 116 Q---AAVILLTDGHNLMGNPVEVKSIYQTPNVCFFHVSFA----- 155
Db 402 QLDGSEVLLITDGBDNTSSCIDEVK---QSGALVHFTALGRAADENAVIEMSKITGSGHF 458
Oy 156 ---DDAEGKAIIDQIVAINSGS 174
Db 459 YVSDEAQNNGILDAFGALTSGN 480

RESULT 8
AAB65272
ID AAB65272 standard; Protein; 919 AA.
XX
AC AAB65272;
XX
DT 02-APR-2001 (first entry)
XX
DE Human PRO1124 (UNQ562) protein sequence SEQ ID NO:379.
XX
KW Human; secreted and transmembrane protein; PRO; cytosolic;
KW cell death; cancer; chromosomal mapping; gene mapping; tissue typing;
KW diagnostic assay.
XX
OS Homo sapiens.
XX
PN MO200073454-A1.
XX
PD 07-DEC-2000.
XX
PF 30-MAR-2000; 2000WO-US08439.
XX
PR 02-JUN-1999; 99WO-US12252.
PR 23-JUN-1999; 99US-0141037.
PR 07-JUL-1999; 99US-0143046.
PR 20-JUL-1999; 99US-0144758.
PR 26-JUL-1999; 99US-0145698.
PR 28-JUL-1999; 99US-0146222.
PR 17-AUG-1999; 99US-0149366.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 08-OCT-1999; 99US-0158663.
PR 30-NOV-1999; 99WO-US28313.
PR 01-DEC-1999; 99WO-US28301.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00376.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US04914.
PR 24-FEB-2000; 2000WO-US05004.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 20-MAR-2000; 2000WO-US07377.
XX
PA (GETH ) GENENTECH INC.
XX
PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
PI Ferrara N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi CJ, Gurney AL, Kljavin IJ, Napier MA, Pan J, Paoni NF;
PI Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WI;
PI Zhang Z;
XX
MPI; 2001-032160/04.

```

DR N-PSDB; AAF44241.

PT PRO polynucleotides used to produce polypeptides used to target

PT bioactive molecules such as toxins, radiolabels or antibodies, to

PT specific cells, to cause targeted cell death -

XX

PS Claim 12; Fig 274; 935pp; English.

XX

CC The present invention describes human secreted and transmembrane PRO

CC proteins. The PRO proteins have cytototoxic activity. The PRO proteins

CC can be used for targeted delivery of bioactive molecules, such as

CC toxins, radiolabels or antibodies, that cause cell death. PRO nucleotide

CC sequences, and their fragments, can be used as hybridisation probes, in

CC chromosomal and gene mapping, and in the generation of anti-sense RNA

CC and DNA. They may also be used to produce transgenic animals which are

CC used to develop and screen therapeutically useful reagents. The PRO

CC nucleotide and protein sequence can be used for tissue typing and in

CC treating cancer. Anti-PRO antibodies can be used in diagnostic assays.

CC AAF44270 to AAF44470 represent PCR primers and hybridisation probes used

CC in the isolation of human PRO sequences. AAF44087 to AAF44269 and

CC AAB65154 to AAB65300 represent human PRO polynucleotide and protein

CC sequences given in the exemplification of the present invention.

XX

SO Sequence 919 AA;

Query Match 8.2%; Score 91.5; DB 22; Length 919;

Best Local Similarity 22.8%; Pred. No. 1.1;

Matches 46; Conservative 26; Mismatches 55; Indels 75; Gaps 8;

QY 17 LVDYSGMMKGVAVREPKIELAKIKNAAMPKXSYOGGLTFAPYSVITPOGSGWNS 76

DB 310 VLDSKSGMG-----GKDRLRNMNQAKHFLQ-----TVENGSMVG 345

QY 77 CV-----AECVNTIKSDLEIFGRLPVG-----DGKMEHYINQMP 115

DB 346 MVHFDSTATVINKLIQIKSSDERNTLMAGLPY-----PLGTSIGSGIKYAFVIGELHS 401

QY 116 Q-----AAVILLTDGHNNIGMNPVEEYKSIYQTNPNVCFFVNSFA----- 155

DB 402 QLDSEVLLTLTDEGSDNTASSCIDEVK---QSGAIVHFYALGRADEAVIEMSKITGSHF 458

QY 156 ---DDEBKAAIIDQIVALNSGS 174

DB 459 VYSDAQNNGLIDAFALTSGN 480

RESULT 9

ABBI1973

ID ABBI1973 standard; peptide; 931 AA.

XX

AC ABBI1973;

XX

DT 11-JAN-2002 (first entry)

XX

DE Human membrane-bound protein homologue, SEQ ID NO:2343.

XX

XX Human; cytokine; cell proliferation; cell differentiation; growth factor;

XX haematopoietic regulation; tissue growth; immunomodulator; activin;

XX inhibin; chemokinesis; chemokinesis; thrombolysis; oncogenesis;

XX proliferation; metastasis; cancer; tumour; haematopoietic disorder;

XX myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;

XX chronic inflammatory condition; proliferative retinopathy;

XX atherosclerosis; coronary heart disease; arterial ischaemia;

XX bone disorder; osteoporosis; vascular growth disorder;

XX tissue regeneration; wound healing; infection; immune disorder;

XX cell culture; drug screening; gene therapy; antiinflammatory;

XX antisthmatic; antiarthritic; haemostatic; antitextosclerotic;

XX cytopathic; osteopathic; vasotropic; cardiant; vitruide; antibacterial;

XX antifungal; vulnery; antitumor.

XX

OS Homo sapiens.

XX

PN W0200157188-A2.

XX

PD 09-AUG-2001.

XX

PF 05-FEB-2001; 2001WO-US03800.

XX

PR 03-FEB-2000; 2000US-0496914.

PR 27-APR-2000; 2000US-0560875.

XX

XX (HYSE-) HYSEQ INC.

XX

PI Tang YT, Liu C, Drmanac RT;

XX

DR WPI; 2001-457740/49.

DR N-PSDB; ABA09217.

XX

PT Human proteins and DNA encoding sequences useful for preventing,

PT treating or ameliorating a medical condition in a mammalian subject

PT e.g. arthritis and cancer -

XX

PS Claim 20; Page 289; 1963pp; English.

XX

CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and

CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The

CC invention also relates to vectors and recombinant host cells comprising a

CC nucleotide of the invention, methods of producing the novel polypeptides,

CC antibodies against the polypeptides, methods of detecting the nucleotides

CC or polypeptides in a sample, and methods of identifying compounds which

CC bind to polypeptides of the invention. Although novel, many of the

CC polypeptides of the invention have homology to known proteins, thereby

CC giving an insight into their probable biological activities, and hence

CC potential therapeutic applications. The polypeptides of the invention may

CC have various activities, including cytokine, cell proliferation or cell

CC differentiation regulatory activity; stem cell growth factor activity;

CC haematopoietic regulatory activity; tissue growth factor activity;

CC immunomodulatory activity; activin or inhibin-related activities;

CC chemotactic or chemokinetic activities; haemostatic, thrombotic or

CC thrombolytic activities; receptor or ligand activities; or may be

CC involved in oncogenesis, cancer cell proliferation or metastasis.

CC Depending on their biological activities, polypeptides and nucleotides of

CC the invention are useful for preventing, treating or ameliorating medical

CC conditions, e.g., by protein or gene therapy. Such conditions include

CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell

CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),

CC proliferative retinopathy, atherosclerosis, coronary heart disease,

CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal

CC vascular growth. Polypeptides involved with tissue regeneration and

CC repair (or nucleic acids encoding them) may be used to promote wound

CC healing (e.g., of burns, incisions and ulcers), while those with

CC immunomodulatory activities may be used in the treatment of viral,

CC bacterial and fungal infections in addition to immune disorders.

CC Polypeptides with growth factor activity may be used in cell cultures to

CC promote cell growth. For example, such polypeptides may be used to

CC manipulate stem cells in culture to give rise to neuroepithelial cells

CC that can be used to augment or replace cells damaged by illness,

CC autoimmune disease or accidental damage. The polypeptides and nucleotides

CC may also be used in the diagnosis of the above conditions, and in drug

CC screening techniques. The present sequence represents a novel human

CC polypeptide of the invention.

XX

SO Sequence 931 AA;

Query Match 8.2%; Score 91.5; DB 22; Length 931;

Best Local Similarity 22.8%; Pred. No. 1.1;

Matches 46; Conservative 26; Mismatches 55; Indels 75; Gaps 8;

QY 17 LVDYSGMMKGVAVREPKIELAKIKNAAMPKXSYOGGLTFAPYSVITPOGSGWNS 76

DB 322 VLDSKSGMG-----GKDRLRNMNQAKHFLQ-----TVENGSMVG 357

QY 77 CV-----AECVNTIKSDLEIFGRLPVG-----DGKMEHYINQMP 115

DB 358 MVHFDSTATVINKLIQIKSSDERNTLMAGLPY-----PLGTSIGSGIKYAFVIGELHS 413

Oy 116 Q---AAVILITPGHNHNLGNMPVEYKSYIQTNPNCVFHVSR----- 155  
Db 414 QLDGSEVLTLTGEDNTASTCIDEVK--QSGAIVHFIALGRAADAVIEMSKITGCSHF 470  
Oy 156 ---DDAEKGAIIDQIVALNSGS 174  
| : | : | : | : | : | :  
| : | : | : | : | : | :  
Db 471 YSDERAGNNGLIDAFGLATSGN 492

RESULT 10  
ID AAU52280  
AAU52280 standard; Protein; 375 AA.  
AC AAU52280;  
XX  
DT 27-FEB-2002 (first entry)  
XX  
DE Propionibacterium acnes immunogenic protein #13176.  
KW SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;  
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
KW dermatological; osteopathic; neuroprotectant.  
XX  
OS Propionibacterium acnes.  
PN WO200181581-A2.  
XX  
PD 01-NOV-2001.  
XX  
PF 20-APR-2001; 2001WO-US12865.  
XX  
PR 21-APR-2000; 2000US-199047P.  
PR 02-JUN-2000; 2000US-208841P.  
XX 07-JUL-2000; 2000US-216747P.  
XX  
PA (CORI-) CORIXA CORP.  
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhacia A;  
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;  
DR N-PsDB; AAS59553.  
DR WPI; 2001-616774/71.  
PS Example 1; SEQ ID NO 13475; 10692p; English.XX  
PT Sequences AAU93105-AAU68017 represent Propionibacterium acnes immunogenic  
PT polypeptides. The proteins and their associated DNA sequences are used in  
PT the treatment, prevention and diagnosis of medical conditions caused by  
PT P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
PT pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.  
XX P. acnes is also involved in infections of bone, joints and the central  
CC nervous system, however it is particularly involved in the inflammatory  
CC lesions associated with acne vulgaris. A method for detecting the  
CC presence or absence of P. acnes in a patient comprises contacting a  
CC sample with a binding agent that binds to the proteins of the invention  
CC and determining the amount of bound protein in the sample. The  
CC polypeptides may be used as antigens in the production of antibodies  
CC specific for P. acnes proteins. These antibodies can be used to  
CC downregulate expression and activity of P. acnes polypeptides and  
CC therefore treat P. acnes infections. The antibodies may also be used as  
CC diagnostic agents for determining P. acnes presence, for example, by  
CC enzyme linked immunosorbent assay (ELISA).  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 375 AA;

```

Query Match Similarity      8.1%; Score 90.5; DB 22; Length 375;
Best Local Silarity        27.3%; Pred. No. 0.37; Indels    21; Gaps
Matches     36; Conservative   23; Mismatches              52;

QY      18 VDSGSMKKHVAVEEP-KTELAKKALIKINAMPKMSYGGLTYFPAPYS-VIIPDGSWN 75
       :|::||:::||::||::||::||::||::||::||::||::||::||::||::||:
Db      150 IIVSSSAMVATDV---EPRSLSAKAARAADFLGDLP-RNNVLVFPASAGVVVP----- 200
       ||::||::||::||::||::||::||::||::||::||::||::||::||::||:
QY      76 SCVAECANVTIKSDLEIFGRLLTPVDGDGMHTETVINQMF-----POAAVILITDG 126
       ::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db      201 PTTDDRAAVASTAIIINQLVPS-FAIGGIYSSINALKLVPDPDKPGQKPAPAIVLISDA 259
       ||::||::||::||::||::||::||::||::||::||::||::||::||:
QY      127 NNLGAMPVEEVK 138
       |:|::||::||::||:
Db      260 TTVGRPRLPAK 271

RESULT 11
ABB48258
ID ABB48258 standard; Protein; 1091 AA.
AC
ABBA48258;
DT 05-FEB-2002 (first entry)
DE Listeria monocytogenes protein #962.
KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KM vitamin B12; bacterial infection; disease.
OS Listeria monocytogenes.
XX
PN WO200177335-A2.
PD 18-OCT-2001.
PF 11-APR-2001; 2001WO-FR01118.
PR 11-APR-2000; 2000FR-0004629.
PX (INSP ) INST PASTEUR.
PI Buchrieser C, Frangeul L, Couve E, Ruesnick C, Feini H, Dehoux P,
PI Dussurget O, Cherouani F, Nedjati H, Glaeer P, Kunst F, Cossart P,
PI Daniels J, Goebel W, Kreft U, Kuhn M, Ng E, Vazquez-Boland JA,
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A,
PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L,
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N,
PI Madueno E, De Pablos B, Wehlund J, Kaerst U, Entian K, Haut J,
PI Rose M, Voss H;
DR MPI; 2002-010914/O1.
PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
PT and prevention of Listeria and related bacterial infections, and
PT related polypeptides -
XX
PS Claim 6; SEQ ID NO 963; 192pp; French.
XX
XX The present invention relates to the genome sequence of Listeria
CC monocytoenes EGD-e (see ABA03041). The genome sequence and fragments of
CC it are useful for selecting probes and primers for detecting genes in L.
CC monocytoenes and related organisms, and for studying genetic
CC polymorphisms and other genomes. The present sequence is a protein
CC encoded by the genome sequence of the present invention. Proteins
CC expressed from the genome sequence are useful for raising specific
CC antibodies, identification of L. monocytoenes and related organisms, and
CC for bioynthesiss and biodegradaton, especially biosynthesis of Vitamin
CC B12. The genome sequence and proteins encoded by it are also useful for
CC selecting compounds that regulate gene expression and cell replication
CC and modulate L. monocytoenes-related diseases. In addition, the genome
CC sequence and proteins encoded by it are useful in pharmaceutical and
CC vaccines compositions for the treatment or prevention of infections by L.
```

CC monocytogenesis and related organisms.

CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences).

XX Sequence 1091 AA;

Query Match 8.0%; Score 90; DB 23; Length 1091;

Best Local Similarity 22.0%; Pred. No. 2.1;  
Matches 56; Conservative 37; Mismatches 74; Indels 88; Gaps 15;

QY 21 SGMAMKAVAREPKIELAKAIIKINAMPMKSYQ--GGITFPAPY-----SVIIPO 71

DB 743 NGKSMRQWADYPE-----DYNARDDEQYMGFDLLVAPRYOEGTEKXVYIPE 793

QY 72 GSNMCAVECAVNTIKSDLEIF--GRITPVGDIKHEETVINOMP--POAAVIL--LTDG 125

DB 794 GEM-----VDIMNGGVHFGERTISTYADV--DILPVFAKAGIIPMMMTDG 837

QY 126 HNNIGMPVEEKSIVQTNPVVCFHV-----VSFADAEQ-----KAI 163

DB 838 Y-QGGVNVGNDLKSIV---DILTFRVYPSGDSSEVSFYDDVNGEMRDISVEDFANEKVS 892

QY 164 ID-----QIVALNSGVLDGLQLQ-----NPAVCEFNVSV---FCQECI 202

DB 893 VDLPMADETTMQVFTEPTSVITIDGADVAKADTLDAFNEATGYGYDTYVNLTVYKAA 952

QY 203 LVTEEVVVLKGVNFA 217

DB 953 KDAKQAIYLVNGVNH 967

RESULT 12

ABB09706 ID ABB09706 standard; Protein: 932 AA.

AC ABB09706;

DT 11-JUN-2002 (first entry)

XX Sequence of H4P heavy chain of inter-alpha-inhibitor protein.

DE MAY-1; blood transfusion; immunosuppressive; mixed lymphocyte reaction;

KW MUR; autoimmune disease; allergic disease; organ rejection;

KW organ transplantation; rheumatism; psoriasis; bronchial asthma;

KW allergic rhinitis; allergic dermatitis; pollinosis; H4P heavy chain;

KW inter-alpha-inhibitor protein.

XX Rattus sp.

OS WO200212495-A1.

PN 14-FEB-2002.

PD 01-AUG-2001; 2001WO-JP06620.

PP 09-AUG-2000; 2000JP-0241169.

PR (MARU-) MARUHO KK.

PA Uchida H, Tanaka H, Kitch Y, Fujimura A, Kobayashi E;

PI WPI, 2002-217191/27.

DR N-PSDB; ABL41969.

XX Rat or human protein MAY-1 induced by homogeneous blood transfusion and

PT exhibiting immunosuppressive activity in a homogeneous mixed lymphocyte

PT reaction, for providing drug compositions to treat e.g. autoimmune

PT diseases -

CC inter-alpha-inhibitor protein. The specification describes MAY-1  
CC protein, which is induced by homogeneous blood transfusion. MAY-1  
CC exhibits an immunosuppressive activity in a homogeneous mixed  
CC lymphocyte reaction (MLR). The MAY-1 protein can formulated into drug  
CC compositions for the prevention or treatment of autoimmune diseases,  
CC allergic diseases, or rejection reaction during organ transplantation,  
CC e.g. Rheumatism, psoriasis, bronchial asthma, allergic rhinitis,  
CC allergic dermatitis and pollinosis.

XX Sequence 932 AA;

Query Match 7.8%; Score 87.5; DB 23; Length 932;

Best Local Similarity 22.1%; Pred. No. 3.1;  
Matches 33; Conservative 31; Mismatches 40; Indels 45; Gaps 9;

QY 16 YLVYSGSMKAVAREPKIELAKAIIKINAMPMKSYGGYTFAPYSVIIPOG--- 72

DB 276 FVIDKSGSMACK-----KIQTREALIKI--LKDISTD-----QFNIIVPGEAN 319

QY 73 SNMCAVECAVNTIKSDLEIFGRITPVGDIKHEETVI-----NQ--MPPQAA-- 118

DB 320 QMEGLVQATEENINRAVDYASKI--PAQGGTNINKAVLSAVELDKSQALPLSKSVSL 378

QY 119 VILITDGHNNIGMPVEEKSIVQTNPV 147

DB 379 ILLTLDGEPTVG-----ETNPXI 396

RESULT 13

ABB09707 ID ABB09707 standard; Protein: 933 AA.

AC ABB09707;

DT 11-JUN-2002 (first entry)

XX Sequence of H4P heavy chain of inter-alpha-inhibitor protein.

DE MAY-1; blood transfusion; immunosuppressive; mixed lymphocyte reaction;

KW MUR; autoimmune disease; allergic disease; organ rejection;

KW organ transplantation; rheumatism; psoriasis; bronchial asthma;

KW allergic rhinitis; allergic dermatitis; pollinosis; H4P heavy chain;

KW inter-alpha-inhibitor protein.

XX Rattus sp.

OS Key Location/Qualifiers

FT Misc-difference 96 /note= "Ieu encoded by AAA"

FT Misc-difference 106 /note= "Tyr encoded by ACT"

PN WO200212495-A1.

PD 14-FEB-2002.

PP 01-AUG-2001; 2001WO-JP06620.

PR 09-AUG-2000; 2000JP-0241169.

PA (MARU-) MARUHO KK.

PI Uchida H, Tanaka H, Kitch Y, Fujimura A, Kobayashi E;

DR WPI, 2002-217191/27.

DR N-PSDB; ABL41970.

XX Rat or human protein MAY-1 induced by homogeneous blood transfusion and

PT exhibiting immunosuppressive activity in a homogeneous mixed lymphocyte

PT reaction, for providing drug compositions to treat e.g. autoimmune

PT diseases -

PS Claim 3; Page 49-53; 85pp; Japanese.

XX	The present sequence encodes the rat H4P heavy chain of
CC	inter-alpha-inhibitor protein. The specification describes MAY-1
CC	protein, which is induced by homogeneous blood transfusion. MAY-1
CC	exhibits an immunosuppressive activity in a homogeneous mixed
CC	lymphocyte reaction (MLR). The MAY-1 protein can formulated into drug
CC	compositions for the prevention or treatment of autoimmune diseases,
CC	allergic diseases, or rejection reaction during organ transplantation,
CC	e.g. pneumatism, psoriasis, bronchial asthma, allergic rhinitis,
CC	allergic dermatitis and pollinosis.
SQ	Sequence 933 AA;
Query Match:	7.8%; Score 87.5; DB 23; Length 933;
Best Local Similarity	22.1%; Pred. No. 3.1;
Matches 33; Conservative 31; Mismatches 40; Indels 45; Gaps 9,	
Oy	16 YLVADVSGSMHMHGVAVREPKRIELAKEAIKINAMPMKMSYGGLYFAPYSVIIPOG--- 72 ::          :           :           :           :
Db	277 FVIDSGSNAGK-----KIOTRALTKI---LKDSLTD-----QFNIIIVSGEAN 320 :           :           :           :           :
Oy	73 SWNSCVAECAVNTIKSIDLEIFGLRPVSDGIKMHTVI-----NG--MPQAA-- 118 :           :           :           :           :
Db	321 QMEQLLVQTTEENLNRAVDYASKI -PAQGCTNINKAVLSAVELLDRKSQAELLPSKSVEL 379 :           :           :           :           :
Oy	119 VILLIDGHNNLGMNPVEEVKSIYQTNPNV 147 :           :           :           :           :
Db	380 ILLTLDGEPTVG-----ETMFKI 397 :           :           :           :           :
RESULT 14	
AABO1830	
ID AABO1830 standard; Protein; 1222 AA.	
XX AC AABO1830;	
XX DT 11-SEP-2000 (first entry)	
DE H. influenzae strain KI mature full-length HMWIA protein, SEQ ID NO:37.	
KM Mature HMW protein; hmw gene; hmwA1; hmwA2; high molecular weight; non-typable Haemophilus influenzae; NTH1; non-encapsulated; recombinant production; Escherichia coli; antibacterial; vaccine; human disease; otitis media; epiglottitis; pneumonia; tracheobronchitis; detection; diagnosis.	
OS Haemophilus influenzae strain KI.	
FH Key Location/Qualifiers	
FT Misc-difference 307 /note= "Encoded by GC"	
XX WO200020609-A2.	
PX PD 13-APR-2000.	
PF 07-OCT-1999; 99WO-CA00938.	
PR 07-OCT-1998; 98US-0167568.	
PR 08-DEC-1998; 98US-0206942.	
PA (CONN-) CONNAUGHT LAB LTD.	
PI Loosmore SM, Yang Y, Klein MH;	
DR WPI; 2000-303789/26.	
N-PsDB; AAA52180.	
PT Nucleic acid molecule for producing recombinant high molecular weight proteins of Haemophilus which are used as a vaccine to provide protection against Haemophilus induced diseases in humans - Claim 8; Fig 20A-R; 307p; English.	

XX		The invention relates to the recombinant production of Haemophilus
CC		influenzae high molecular weight (HMW) proteins in Escherichia coli. The
CC		expression construct used to effect recombinant expression comprises a
CC		promoter functional in E. coli (e.g., the tr promoter) operably linked
CC		to a modified hmwABC operon from a non-typable (non-encapsulated) H.
CC		influenzae (NTHI). Most HMW-expressing NTHI strains contain two hmw gene
CC		clusters termed hmwLABC and hmwZABC. Each hmwABC operon comprises hmwA,
CC		hmwB and hmwC genes. The hmwA genes encode the structural HMWA proteins
CC		and the hmwB and hmwC genes encode accessory proteins which are
CC		responsible for post-translational processing and secretion of the HMWA
CC		proteins. The modified hmwABC operon used in the expression construct of
CC		the invention contains an A gene modified such that it encodes only the
CC		mature HMWA. The invention also discloses hmwA genes (AA452175-A52198)
CC		and HMWA proteins (AAB01824-B01849) from the non-typable H. influenzae
CC		strains D9C, KI, KX1, LDCD2, PMH, 15 and 12. The nucleic acids and
CC		vectors are used for the production of recombinant H. influenzae HMW
CC		proteins which can be used as vaccines to mediate a humoral or
CC		cell-mediated immune response to provide protection against diseases in
CC		humans caused by H. influenzae (e.g., otitis media, epiglottitis,
CC		pneumonia and tracheobronchitis). The HMW proteins are also useful as
CC		antigens in immunoassays for detecting antibodies against Haemophilus,
CC		HMW proteins and/or HMW peptides. The nucleotide sequences encoding the
CC		HMW proteins can be used to isolate and clone hmw genes from other
CC		non-typable strains of Haemophilus via hybridisation reactions. The
CC		present sequence represents a mature HMWA protein from a non-typable
CC		strain of H. influenzae.
SQ	Sequence	1222 AA;
XX		
XX		
Query Match	7.7%	Score 86; DB 21; Length 1222;
Best Local Similarity	17.8%	Pred. No. 6.8;
Matches	48; Conservative	38; Mismatches 98; Indels 86; Gaps 7
Oy	11 VESVYLVYVSGSMKKHYAVREPKIELAKAIL-----KINAM 50	
Db	329 ISSQNPSASGGSSLKFKSKSGTHAAFTINDILNATGNSLNGVAGIDSNLKKSLIAN 388	
Oy	51 PKMSVOGLYTFA-----PYSVITPGGSWNCSVAECNAVTTIKSDLEIFGRLPVGVG 102	
Db	389 KNITEFGGNITLAADKPRIEIKNTIVKEGA-NVLIRSANYNDSALISIRNVTKGN- 446	
Oy	103 IMHEEVINQMPPQAVALITLDGHNNLGNNPVEVKYSIQTPNVCFFHVSPADD----- 157	
Db	447 -----LTVTGSAINIEKNLTVESAKFLANPNVSFNVSGLPFDNGSKN 489	
Oy	158 ---AEKAIIDDI VALNCGSVLV-----DGQLLON 185	
Db	490 ISIAKGAFLFKIE-NTGSLNITTTSKDSNHHTIIKGNITNRKGDLINTNGDNTEIQIG 547	
Oy	186 PAVCGEFNVSHFCOEQILTVEEVVLRGVN 215	
Db	548 GNISOKEGNLITISDKVITERITTIKAVGN 577	
RESULT 15		
AAB01828	ID	AAB01828 standard; Protein; 1228 AA.
XX		
XX	AAB01828;	
XX		
DT	11-SEP-2000	(first entry)
XX		
DE	Haemophilus influenzae strain KI full-length HMWA protein, SEQ ID NO:34.	
XX		
KM	HMW protein; hmw gene; hmwA1; hmwA2; high molecular weight;	
KM	non-typable Haemophilus influenzae; NTHI; non-encapsulated;	
KM	recombinant production; Escherichia coli; antibacterial; vaccine;	
KM	human disease; otitis media; epiglottitis; pneumonia; tracheobronchitis;	
KM	detection; diagnosis.	
XX		
XX	Haemophilus influenzae strain KI.	
XX		

Key Location/Qualifiers  
 Misc-difference 313 /note= "Encoded by GG"  
 FT  
 PN  
 XX  
 MO200020609-A2.  
 PD  
 13-APR-2000.  
 PF  
 07-OCT-1999; 99MO-CA00938.  
 PR  
 07-OCT-1998; 98US-0167568.  
 PR  
 08-DEC-1998; 98US-0206942.  
 PA  
 (CONN-) CONNAUGHT LAB LTD.  
 PI  
 Loosmore SM, Yang Y, Klein MH;  
 DR  
 WPI; 2000-303789/26.  
 DR  
 N-PSDB; AAA52179.  
 XX  
 PS  
 Claim 12; Fig 20A-R; 307pp; English.  
 XX  
 CC  
 The invention relates to the recombinant production of Haemophilus  
 CC  
 influenzae high molecular weight (HMW) proteins in Escherichia coli. The  
 CC  
 expression construct used to effect recombinant expression comprises a  
 CC  
 promoter functional in E. coli (e.g., the T7 promoter) operably linked  
 CC  
 to a modified hmwaB operon from a non-typeable (non-encapsulated) H.  
 CC  
 influenzae (NTHI). Most HMW-expressing NTHI strains contain two hmwa  
 CC  
 clusters termed hmwa1ABC and hmwa2ABC. Each hmwaB operon comprises hmwa,  
 CC  
 and the hmwb and hmwC genes encode accessory proteins which are  
 CC  
 responsible for post-translational processing and secretion of the HMWA  
 CC  
 proteins. The modified hmwaB operon used in the expression construct of  
 CC  
 the invention contains an A gene modified such that it encodes only the  
 CC  
 mature HMWA. The invention also discloses hmwa genes (AAA52175-A52198)  
 CC  
 and HMWA proteins (AAB01824-B01849) from the non-typeable H. influenzae  
 CC  
 strains J09c, K1, K21, LDC2, PMH1, 15 and 12. The nucleic acids and  
 CC  
 vectors are used for the production of recombinant H. influenzae HMW  
 CC  
 proteins which can be used as vaccines to mediate a humoral or  
 CC  
 cell-mediated immune response to provide protection against diseases in  
 CC  
 humans caused by H. influenzae (e.g., otitis media, epiglottitis,  
 CC  
 pneumonia and tracheobronchitis). The HMW proteins are also useful as  
 CC  
 antigens in immunoassays for detecting antibodies against Haemophilus,  
 CC  
 HMW proteins and/or HMW peptides. The nucleotide sequences encoding the  
 CC  
 HMW proteins can be used to isolate and clone hmwa genes from other  
 CC  
 non-typeable strains of Haemophilus via hybridisation reactions. The  
 CC  
 present sequence represents an HMWA protein from a non-typeable strain of  
 CC  
 H. influenzae.  
 CC  
 CC  
 SQ  
 Sequence 1228 AA;  
 Query Match 7.7%; Score 86; DB 21; Length 1228;  
 Best Local Similarity 17.8%; Pred. No. 6.8;  
 Matches 48; Conservative 38; Mismatches 98; Indels 86; Gaps 7;  
 QY 11 VESYNYLVDSGSMKGVAVREPKIELAKEAIL-----KINAA 50  
 DB 335 IESQNFSAAGSSSLKFKSEGSTHAAFTIKNDLILNATGNSISLQVAGIDSNLKKSLIAN 394  
 QY 51 PKMSYOGGLYTF-----PYSVILPQSGMSNCVACAVNTIKSDLEIFGRRLTPVGDG 102  
 DB 395 KNITFEGGNITLADKKPIETIKGNITVEKA-NTLRSANYGNDKSAISIRGNVNTKGN- 452  
 QY 103 IKMHEITVINCMPQAAVILITDGHNNILGMNPEEVKSIYQTNPNVCFHVSFADD----- 157  
 DB 453 -----LTVGSAINIEKNLTVGSAKFLANPNYSFVNSGLFNDQKSN 495  
 QY 158 ---AEGKAIIDQIVANLSGSVIV-----DGLQLLQN 185

DB 496 ISIAKGAIFPDIE--NTGSLNITTKSDSNHHTIIKGNITNRKGDNLITNGDNTETIQIG 553  
 QY 186 PAVCOEFVNSVFCQEQILVTEEVVLRGVN 215  
 DB 554 GNISQKEGNLTSSDKNITRITITIKAGVN 583

Search completed: April 9, 2003, 14:30:08  
 Job time : 79 secs





GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: April 9, 2003, 14:29:23 ; Search time 28 Seconds  
(without alignments)  
229.078 Million cell updates/sec

Title: US-10-034-500-2  
Perfect score: 1119  
Sequence: 1 AEVTASCTKRVESYNYLVY.....QEQILVTEVVLGVNFAR 218

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/Backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	92	8.2	885	US-09-074-579-5	Sequence 5, Appli
2	92	8.2	885	US-09-388-774-5	Sequence 5, Appli
3	91.5	8.2	917	US-09-049-698-41	Sequence 41, Appli
4	86	7.7	915	US-09-206-942-35	Sequence 35, Appli
5	86	7.7	1222	US-09-206-942-37	Sequence 37, Appli
6	86	7.7	1228	US-09-206-942-34	Sequence 34, Appli
7	80	7.1	391	US-08-968-563-11	Sequence 11, Appli
8	80	7.1	391	US-08-969-683A-11	Sequence 11, Appli
9	80	7.1	391	US-09-297-928-7	Sequence 7, Appli
10	79.5	7.1	1095	US-09-206-942-45	Sequence 45, Appli
11	79.5	7.1	1101	US-09-206-942-43	Sequence 43, Appli
12	79	7.1	141	US-08-470-179-17	Sequence 17, Appli
13	79	7.1	141	US-08-470-179-19	Sequence 19, Appli
14	79	7.1	2183	US-08-348-891A-7	Sequence 7, Appli
15	79	7.1	2183	US-08-905-817-7	Sequence 7, Appli
16	77.5	6.9	317	US-09-109-205-18	Sequence 18, Appli
17	77.5	6.9	317	US-08-940-424-5	Sequence 5, Appli
18	77.5	6.9	903	US-09-193-562D-46	Sequence 46, Appli
19	77	6.9	141	US-08-470-179-18	Sequence 18, Appli
20	77	6.9	141	US-08-470-179-20	Sequence 20, Appli
21	76.5	6.8	320	US-08-193-562D-28	Sequence 28, Appli
22	76	6.8	320	US-08-245-511-4	Sequence 4, Appli
23	76	6.8	514	US-08-600-993A-4	Sequence 35, Appli
24	76	6.8	514	US-08-688-988-35	Sequence 5, Appli
25	76	6.8	666	US-08-485-284A-5	Sequence 2, Appli
26	76	6.8	666	US-08-961-083-2	Sequence 6, Appli
27	76	6.8	682	US-08-481-435-6	

28	75.5	6.7	708	US-09-131-648-2	Sequence 2, Appli
29	75.5	6.7	1152	US-08-476-062A-43	Sequence 43, Appli
30	75.5	6.7	1152	PCT-US96-01314-43	Sequence 43, Appli
31	75.5	6.7	1152	5424399-2	Patent No. 5424399
32	75.5	6.7	1153	US-08-173-497-3	Sequence 3, Appli
33	75.5	6.7	1153	US-08-286-889-3	Sequence 3, Appli
34	75.5	6.7	1153	US-08-485-618-3	Sequence 3, Appli
35	75.5	6.7	1153	US-08-362-652-3	Sequence 3, Appli
36	75.5	6.7	1153	US-08-605-672-3	Sequence 3, Appli
37	75.5	6.7	1153	US-08-482-293A-3	Sequence 3, Appli
38	75.5	6.7	1153	US-08-943-363-3	Sequence 3, Appli
39	75.5	6.7	1153	US-09-193-043-3	Sequence 3, Appli
40	75.5	6.7	1153	US-09-688-307A-3	Sequence 3, Appli
41	75.5	6.7	1170	US-08-476-062A-42	Sequence 42, Appli
42	75.5	6.7	1170	PCT-US96-01314-42	Sequence 42, Appli
43	74	6.6	911	US-09-074-579-4	Sequence 4, Appli
44	74	6.6	911	US-09-388-774-4	Sequence 4, Appli
45	74	6.6	953	US-08-500-857A-2	Sequence 2, Appli

## ALIGNMENTS

RESULT 1  
US-09-074-579-5  
Sequence 5, Application US/09074579  
Patent No. 6001596  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Guglier, Karl J.  
TITLE OF INVENTION: GROWTH-ASSOCIATED TRYPsin-TYPE  
NUMBER OF INVENTION: INHIBITOR HEAVY CHAIN PRECURSOR  
CORRESPONDENCE ADDRESS:  
ADDRESS: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 6.1/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/074,579  
FILING DATE: HEREWITH  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Cerrone, Michael C  
REGISTRATION NUMBER: 39,132  
REFERENCE/DOCKET NUMBER: PF-0505 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 885 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENBANK  
CLONE: gi288563  
US-09-074-579-5  
Query Match 8.2%; Score 92; DB 3; Length 885;

Best Local Similarity 24.8%; Pred. No. 0.082;  
Matches 34; Conservative 30; Mismatches 45; Indels 28; Gaps 7;

Qy 16 YLVDSGSMKMHVAVREPKIELAKEALIKINAAMPKMSYOGSLYTFAPYSVIIPOGSMN 75  
Db 283 FVIDISGSMAGR-----KLEOTKEALRLILEDKMEEDY---LNFLIFSGDV--STWK 329  
Qy 76 SCVAECA-----VNTIKSDLEIFGLTPVGDGIKMHETVIN-----QMPPOAA--VIL 121  
Db 330 EHLVQATPENLOEARTFVKSMEDKG-MTNINDGLRGISMLNKAREHRIPERSISYIM 388  
Qy 122 LTDGNNI.GMNPVEEVK 138  
Db 389 LTDGDANVGESRPEKIQ 405

## RESULT 2

US-09-368-774-5  
; Sequence 5, Application US/09388774  
; Patent No. 622891  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Patterson, Chandra  
; TITLE OF INVENTION: GROWTH-ASSOCIATED TRYPsin-TYPE  
; TITLE OF INVENTION: INHIBITOR HEAVY CHAIN PRECURSOR  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 6.1/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/368,774  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/074,579  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ceirone, Michael C  
; REGISTRATION NUMBER: 39,132  
; REFERENCE/DOCKET NUMBER: PF-0505 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-855-0555  
; TELEFAX: 650-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 885 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GENEBANK  
; CLONE: g1288563  
; US-09-368-774-5

Query Match 8.2%; Score 92; DB 4; Length 885;  
Best Local Similarity 24.8%; Pred. No. 0.082;  
Matches 34; Conservative 30; Mismatches 45; Indels 28; Gaps 7;

Qy 16 YLVDSGSMKMHVAVREPKIELAKEALIKINAAMPKMSYOGSLYTFAPYSVIIPOGSMN 75  
Db 283 FVIDISGSMAGR-----KLEOTKEALRLILEDKMEEDY---LNFLIFSGDV--STWK 329

Qy 76 SCVAECA-----VNTIKSDLEIFGLTPVGDGIKMHETVIN-----QMPPOAA--VIL 121  
Db 330 EHLVQATPENLOEARTFVKSMEDKG-MTNINDGLRGISMLNKAREHRIPERSISYIM 388  
Qy 122 LTDGNNI.GMNPVEEVK 138  
Db 389 LTDGDANVGESRPEKIQ 405

## RESULT 3

US-09-049-698-41  
; Sequence 41, Application US/09049698  
; Patent No. 6368792  
; GENERAL INFORMATION:  
; APPLICANT: BILLING-MEDEL, PATRICIA A.  
; APPLICANT: COHEN, MAURICE  
; APPLICANT: COLBITTS, TRACEY L.  
; APPLICANT: FRIEDMAN, PAULA N.  
; APPLICANT: HAYDEN, MARK  
; APPLICANT: KLAS, MICHAEL R.  
; APPLICANT: ROBERTS-RAPP, LISA  
; APPLICANT: RUSSELL, JOHN C.  
; APPLICANT: STROUPE, STEPHEN D.  
; TITLE OF INVENTION: REAGENTS AND METHODS FOR THE  
; TITLE OF INVENTION: USEFUL FOR DETECTING DISEASES OF THE GASTROINTESTINAL  
; TITLE OF INVENTION: TRACT  
; NUMBER OF SEQUENCES: 51  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Abbott Laboratories  
; STREET: 100 Abbott Park Road  
; CITY: Abbott Park  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 60064-3500  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/049,698  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/828,856  
; FILING DATE: 31-MAR-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Becker, Cheryl L.  
; REGISTRATION NUMBER: 35,441  
; REFERENCE/DOCKET NUMBER: 6068 US.P1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 847/935-1729  
; TELEFAX: 847/938-2623  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 41:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 917 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: No. 6368792e  
; US-09-049-698-41

Query Match 8.2%; Score 91.5; DB 4; Length 917;  
Best Local Similarity 22.8%; Pred. No. 0.099;  
Matches 46; Conservative 26; Mismatches 55; Indels 75; Gaps 8;

Qy 17 YLVDSGSMKMHVAVREPKIELAKEALIKINAAMPKMSYOGSLYTFAPYSVIIPOGSMN 76  
Db 310 YLDKSGSMG-----GKDRINRMNOAAHFLIQ-----TVENGSWVG 345  
Qy 77 CV-----AECAVNTIKSDLEIFGLTPVG-----DGIMHETVINQMP 115

```

Db 346 MWHDSATATVTKIQIKSSDERNTLMAGLPTY----PLGTSICSGIKYAFQYIGELHS 401
Qy 116 Q---AAVILLTDGHNNGMNPVEEVKSIYQTNPNVCFHVVSFA----- 155
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 402 QLDGSEVLLTLDGSDNTASCSIDEVK---QSGALVHFIALGRADEAVTEMSKITGSHF 458
Qy 156 ----DDAGKAIIDQIVALNSGS 174
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 459 YVSDAQNNGLIDAFGALTSGN 480

```

# RESULT 4

```

US-09-206-942-35
; Sequence 35, Application US/09206942
; Patent No. 6432669
; GENERAL INFORMATION:
; APPLICANT: Loomore, Sheena M.
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: Protective Recombinant Haemophilus influenzae High
; FILE REFERENCE: 1038-861 MIS:jb
; CURRENT APPLICATION NUMBER: US/09/206,942
; EARLIER FILING DATE: 1998-12-08
; EARLIER APPLICATION NUMBER: 09/167,568
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 35
; LENGTH: 915
; TYPE: PR1
; ORGANISM: Haemophilus influenzae
US-09-206-942-35

```

```

Query Match 7.7%; Score 86; DB 4; Length 915;
Best Local Similarity 17.8%; Pred. No. 0.44;
Matches 48; Conservative 38; Mismatches 98; Indels 86; Gaps 7;

```

```

Qy 11 VESYNYLVDSGSMGMKGVAVREPKIELAKEAIL-----KINAM 50
; : : : : : : : : : : : : : : : : : : : : : : : :
Db 22 IESQNFASGSSSLKFKSEGSTHAFTIKDILLNATGNSISLQVAGIDSNLKKSLIAN 81
; : : : : : : : : : : : : : : : : : : : : : : : :
Qy 51 PKMSYQGLYTPA-----PYSVIIPOGSMNSCVACAVNTIKSDLEIFGRLLTPVGDG 102
; : : : : : : : : : : : : : : : : : : : : : : : :
Db 82 KNIFEGGNITLADKKPIEIKGNTVEGA-NVTLRSANYGNDKSLISIRGNVTNKG- 139
; : : : : : : : : : : : : : : : : : : : : : : : :
Qy 103 IKMHEVYINQMPQAAVILLTDGHNNGMNPVEEVKSIYQTNPNVCFHVVSFAD- 157
; : : : : : : : : : : : : : : : : : : : : : : : :
Db 140 -----LTVGSAINIEKULTVEGSAKFLANPVSFVSGLFDNQGKSN 182
; : : : : : : : : : : : : : : : : : : : : : : : :
Qy 158 ---AEGKAIIDQIVALNSGSVTV-----DGLQLLON 185
; : : : : : : : : : : : : : : : : : : : : : : : :
Db 183 ISIAKGAIFPDIE--NTGSLNITTKSDSNHHTIIKGNITRKGDNLITNNGDTEIQIG 240
; : : : : : : : : : : : : : : : : : : : : : : : :
Qy 186 PAVQGEFVNSVFCQEQILVTEEVVLRGNV 215
; : : : : : : : : : : : : : : : : : : : : : : : :
Db 241 GNISQKEGNLTISDKNITERITIKAGVN 270

```

```

RESULT 5
US-09-206-942-37
; Sequence 37, Application US/09206942
; Patent No. 6432669
; GENERAL INFORMATION:
; APPLICANT: Loomore, Sheena M.
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: Protective Recombinant Haemophilus influenzae High
; FILE REFERENCE: 1038-861 MIS:jb
; CURRENT APPLICATION NUMBER: US/09/206,942
; EARLIER FILING DATE: 1998-12-08
; EARLIER APPLICATION NUMBER: 09/167,568

```

```

; EARLIER FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 37
; LENGTH: 1222
; TYPE: PR1
; ORGANISM: Haemophilus influenzae
US-09-206-942-37

```

```

Query Match 7.7%; Score 86; DB 4; Length 1222;
Best Local Similarity 17.8%; Pred. No. 0.77;
Matches 48; Conservative 38; Mismatches 98; Indels 86; Gaps 7;

```

```

Qy 11 VESYNYLVDSGSMGMKGVAVREPKIELAKEAIL-----KINAM 50
; : : : : : : : : : : : : : : : : : : : : : : : :
Db 329 IESQNFASGSSSLKFKSEGSTHAFTIKDILLNATGNSISLQVAGIDSNLKKSLIAN 388
; : : : : : : : : : : : : : : : : : : : : : : : :
Qy 51 PKMSYQGLYTPA-----PYSVIIPOGSMNSCVACAVNTIKSDLEIFGRLLTPVGDG 102
; : : : : : : : : : : : : : : : : : : : : : : : :
Db 389 KNIFEGGNITLADKKPIEIKGNTVEGA-NVTLRSANYGNDKSLISIRGNVTNKG- 446
; : : : : : : : : : : : : : : : : : : : : : : : :
Qy 103 IKMHEVYINQMPQAAVILLTDGHNNGMNPVEEVKSIYQTNPNVCFHVVSFAD- 157
; : : : : : : : : : : : : : : : : : : : : : : : :
Db 447 -----LTVGSAINIEKULTVEGSAKFLANPVSFVSGLFDNQGKSN 489
; : : : : : : : : : : : : : : : : : : : : : : : :
Qy 158 ---AEGKAIIDQIVALNSGSVTV-----DGLQLLON 185
; : : : : : : : : : : : : : : : : : : : : : : : :
Db 490 ISIAKGAIFPDIE--NTGSLNITTKSDSNHHTIIKGNITRKGDNLITNNGDTEIQIG 547
; : : : : : : : : : : : : : : : : : : : : : : : :
Qy 186 PAVQGEFVNSVFCQEQILVTEEVVLRGNV 215
; : : : : : : : : : : : : : : : : : : : : : : : :
Db 548 GNISQKEGNLTISDKNITERITIKAGVN 577

```

# RESULT 6

```

US-09-206-942-34
; Sequence 34, Application US/09206942
; Patent No. 6432669
; GENERAL INFORMATION:
; APPLICANT: Loomore, Sheena M.
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: Protective Recombinant Haemophilus influenzae High
; FILE REFERENCE: 1038-861 MIS:jb
; CURRENT APPLICATION NUMBER: US/09/206,942
; EARLIER FILING DATE: 1998-12-08
; EARLIER APPLICATION NUMBER: 09/167,568
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 34
; LENGTH: 1228
; TYPE: PR1
; ORGANISM: Haemophilus influenzae
US-09-206-942-34

```

```

Query Match 7.7%; Score 86; DB 4; Length 1228;
Best Local Similarity 17.8%; Pred. No. 0.77;
Matches 48; Conservative 38; Mismatches 98; Indels 86; Gaps 7;

```

```

Qy 11 VESYNYLVDSGSMGMKGVAVREPKIELAKEAIL-----KINAM 50
; : : : : : : : : : : : : : : : : : : : : : : : :
Db 335 IESQNFASGSSSLKFKSEGSTHAFTIKDILLNATGNSISLQVAGIDSNLKKSLIAN 394
; : : : : : : : : : : : : : : : : : : : : : : : :
Qy 51 PKMSYQGLYTPA-----PYSVIIPOGSMNSCVACAVNTIKSDLEIFGRLLTPVGDG 102
; : : : : : : : : : : : : : : : : : : : : : : : :
Db 395 KNIFEGGNITLADKKPIEIKGNTVEGA-NVTLRSANYGNDKSLISIRGNVTNKG- 452
; : : : : : : : : : : : : : : : : : : : : : : : :
Qy 103 IKMHEVYINQMPQAAVILLTDGHNNGMNPVEEVKSIYQTNPNVCFHVVSFAD- 157
; : : : : : : : : : : : : : : : : : : : : : : : :
Db 453 -----LTVGSAINIEKULTVEGSAKFLANPVSFVSGLFDNQGKSN 495

```

Qy 158 ---AEGKAIIDQVIALNSGSLV-----DGLQILON 185  
Db 496 ISIKGKGIPIODIE--NTGSLNITTKSDSNHHTIKGNITNRKGDPLNTNNGDNTEIIG 553  
Qy 186 PAVCOEFVNSVFCOEQILVTEEVVLRGN 215  
Db 554 GNISQKEGNLTISDKVNITERITIKAGVN 583

## RESULT 7

US-08-968-563-11  
; Sequence 11, Application US/08968563  
; Patent No. 6013494  
; GENERAL INFORMATION:  
; APPLICANT: CHARLES E. NAKAMURA  
; APPLICANT: ANTHONY A. GATENBY  
; APPLICANT: AMY (KUANG-HUA) HSU  
; APPLICANT: RICHARD D. LA REAU  
; APPLICANT: SHARON L. HAYNIE  
; APPLICANT: MARIA DIAZ-TORRES  
; APPLICANT: DONALD E. TRIMBUR  
; APPLICANT: GREGORY M. WHITED  
; APPLICANT: VASANTHA NAGARAJAN  
; APPLICANT: MARK S. PAYNE  
; APPLICANT: STEPHEN K. PICCAGGIO  
; APPLICANT: RAMESCH V. NAIR  
; TITLE OF INVENTION: METHOD FOR THE RECOMBINANT  
; TITLE OF INVENTION: PRODUCTION OF 1,3-PROPANEDIOL  
; NUMBER OF SEQUENCES: 49  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY  
; STREET: 1007 MARKET STREET  
; CITY: WILMINGTON  
; STATE: DELAWARE  
; COUNTRY: U.S.A.  
; ZIP: 19898  
; ADDRESSEE: GENENCOR INTERNATIONAL, INC.  
; STREET: 4 CAMBRIDGE PLACE  
; STREET: 1870 SOUTH WINTON ROAD  
; CITY: ROCHESTER  
; STATE: NEW YORK  
; COUNTRY: U.S.A.  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.50 INCH DISKETTE  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95  
; SOFTWARE: MICROSOFT WORD VERSION 7.0A  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/968,563  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/030,601  
; FILING DATE: NOVEMBER 13, 1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: FLOYD, LINDA AXAMETHY  
; REGISTRATION NUMBER: 33,692  
; REFERENCE/DOCKET NUMBER: CR-9982  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 302-892-8112  
; TELEFAX: 302-773-0164  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 391 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; ORIGINAL SOURCE:  
; ORGANISM: GPD1  
; US-08-968-563-11

Query Match 7.1%; Score 80; DB 3; Length 391;

Best Local Similarity 22.1%; Pred. No. 0.6;  
Matches 40; Conservative 27; Mismatches 58; Indels 56; Gaps 10;

Qy 64 PYSV-IIPQGSWNSGVACAVNTIKSDLEIFGRILPVG-----DGLKHETV-----I 110  
Db 34 PFKVTVISGSGWTTIAKVAENCKGYEVAPVQKMWVEEELNGEKLTLITNRHONV 93  
Qy 111 NQMPQAAVILLTDGHNHNGMP--VEEVKSI-----YQTNPNVCFHVASFADDAEG 160  
Db 94 KYLPQ---ITLPD---NLVANPDLIDSVKQVDILIVENIPHQFLPRICSLKGHVD---- 142  
Qy 161 KAIIDQVIALNSGSLVLDGLQILONPAVCOEFVNSVFCOEQILVTEEVV---LRGNV 216  
Db 143 -SHVRAISCLGFEVGAQVQLLS-----YITEILGICGALSGANI 184  
Qy 217 A 217  
Db 185 A 185

## RESULT 8

US-08-969-683A-11  
; Sequence 11, Application US/08969683A  
; Patent No. 6136576  
; GENERAL INFORMATION:  
; APPLICANT: GENENCOR INTERNATIONAL, INC.  
; TITLE OF INVENTION: METHOD FOR THE RECOMBINANT  
; TITLE OF INVENTION: PRODUCTION OF 1,3 PROPANEDIOL  
; NUMBER OF SEQUENCES: 68  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genencor International, Inc.  
; STREET: 4 Cambridge Place  
; STREET: 1870 South Winton road  
; CITY: Rochester  
; STATE: NY  
; COUNTRY: U.S.A.  
; ZIP: 14618  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows  
; SOFTWARE: FastSeq for Windows Version 2.0b  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/969,683A  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US97/20873  
; FILING DATE: 13-NOV-1997  
; APPLICATION NUMBER: 60/030,601  
; FILING DATE: 13-NOV-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Glaister, Debra  
; REGISTRATION NUMBER: 33,888  
; REFERENCE/DOCKET NUMBER: GC 369-2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-864-7620  
; TELEFAX: 650-845-6504  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 391 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; ORIGINAL SOURCE:  
; ORGANISM: GPD1  
; US-08-969-683A-11

Query Match 7.1%; Score 80; DB 4; Length 391;  
Best Local Similarity 22.1%; Pred. No. 0.6;  
Matches 40; Conservative 27; Mismatches 58; Indels 56; Gaps 10;



```

; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: Protective Recombinant Haemophilus Influenzae High
; TITLE OF INVENTION: Molecular Weight Proteins
; FILE REFERENCE: 1038-861 MIS:jdb
; CURRENT APPLICATION NUMBER: US/09/206,942
; CURRENT FILING DATE: 1998-12-08
; EARLIER APPLICATION NUMBER: 09/167,568
; EARLIER FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 43
; LENGTH: 1101
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
; US-09-206-942-43

Query Match          7.1%; Score 79.5; DB 4; Length 1101;
Best Local Similarity 19.4%; Pred. No. 3.5;
Matches 48; Conservative 41; Mismatches 90; Indels 69; Gaps 9;

QY 11 VESVYVLDYSGMMKGVAVREPKIELAKEALIKINAMPKSYOGGLYTFP----- 63
DB 361 IESDITLTAATGNSILNOVAGIGDNQ-----KSLVANKNITTEGGNITTLADKKPIE 413
QY 64 -PYSVILPQGSWNSCVABCAVNTIKSDLEIFGRLLTPVGDGIRKHEVYINQMPPQAAVILL 122
DB 414 IKGNTTVEGA-NVTLRSANVGNDSKLSIRGNVTNKN-----LTV 454
QY 123 TDGHNLLMNPVEVYSIQTPNVCFHVVSFAD-----ARGKAIIDOI----- 167
DB 455 TGSAINIEKNLTVESGAKFLANPNYSFNVSGLPDNGKSNISIAKGAFKIDINNTKSLN 514
QY 168 VALNSG---VLVDG-----LQLLQNP-----AVGQEFNVSFCQEOILVTEE 207
DB 515 ITTMSDSAVRTIIEGNTINSNGDLITTDKNNAEIQIGNISQKGNLTISDKINITIQ 574
QY 208 VVVLGVN 215
DB 575 ITIKGVN 582

RESULT 12
US-08-470-179-17
; Sequence 17, Application US/08470179
; Patent No. 5645994
; GENERAL INFORMATION:
; APPLICANT: Huang Ph.D, Wei Mun
; TITLE OF INVENTION: Method and Compositions for
; TITLE OF INVENTION: Identification of Species in a Sample
; NUMBER OF SEQUENCES: 207
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Trask, Britt and Rossa
; STREET: P.O. Box 2550
; CITY: Salt Lake City
; STATE: Utah
; COUNTRY: USA
; ZIP: 84110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,179
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sweigert Ph.D, Susan E.
; REGISTRATION NUMBER: 36,289
; REFERENCE/DOCKET NUMBER: 2601
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 801-532-1922
; TELEFAX: 801-531-9168
```

```

; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 141 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli
; US-08-470-179-17

Query Match          7.1%; Score 79; DB 1; Length 141;
Best Local Similarity 19.8%; Pred. No. 0.16;
Matches 25; Conservative 26; Mismatches 59; Indels 16; Gaps 4;

QY 65 YSVIIPQGSWNSCVABCAVNTIKSDLEIFGRLLTPVGDGIRKHEVYINQMPPQAAVILLTD 124
DB 12 YAMNVLAGDMNKAKYKSA--RVGD--VIGKYPHGDG-AVYDITVMAQPFSLRYMLVD 66
QY 125 GHNNG-----MNPVEVYSIQTPNVCFHVVSFADAEKAIIDQIVALNSG 173
DB 67 GQGNFGSIDGSAAMRYTEIRLAKIAHELMADLEKETVDFVDNDYGEKIPDVMPTKIP 126
QY 174 SVLVNG 179
DB 127 NLVNG 132

RESULT 13
US-08-470-179-19
; Sequence 19, Application US/08470179
; Patent No. 5645994
; GENERAL INFORMATION:
; APPLICANT: Huang Ph.D, Wei Mun
; TITLE OF INVENTION: Method and Compositions for
; TITLE OF INVENTION: Identification of Species in a Sample
; NUMBER OF SEQUENCES: 207
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Trask, Britt and Rossa
; STREET: P.O. Box 2550
; CITY: Salt Lake City
; STATE: Utah
; COUNTRY: USA
; ZIP: 84110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,179
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sweigert Ph.D, Susan E.
; REGISTRATION NUMBER: 36,289
; REFERENCE/DOCKET NUMBER: 2601
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 801-532-1922
; TELEFAX: 801-531-9168
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 141 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
```





SEQUENCE CHARACTERISTICS:  
LENGTH: 2183 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-905-817-7

Query Match 7.1%; Score 79; DB 2; Length 2183;  
Best Local Similarity 17.7%; Pred. No. 12;  
Matches 54; Conservative 52; Mismatches 81; Indels 118; Gaps 13;

QY 7 CTKRVESYNYLVDSYSGMMKHAIVREPKIELAKEALLKINAMPKMSYQGGLYT----- 61  
Db 1777 CLEPGEDEGLFLGEGSGSMILTY-----KEILKLN---KCFYNSGVSANSRSG 1820  
QY 62 ---PAPYSVII-----PQSGMNSCVAC-----AVNTIK 87  
Db 1821 QRELAPYFSEVGLVEHRMGVGNIVKVLFNGRPEVTWGSV-DCFNFIVSNIFTSSVGFLH 1879  
QY 88 SDLE-----IFGLTPV-----GDGIK-----MHETV 109  
Db 1880 SDIETLPKNDTEKLELALISMLLGLCKIGSILVIKLMPSGDPVQGFISYVGSYYRE 1939  
QY 110 INQMPQ-----AAVILLIDGHNNLGMNPVEEVK-----SIYOTNPVCFHVVSFPADD 157  
Db 1940 VNLVYPRYSNFIESTSYLVMTDLKANRLMNP-EKIKQOIIESSVRTSPGLIGHIILIKOL 1998  
QY 158 AEGKAIIDQIYVALNS-----GSYLVDGLQLQNPAYCOEFVNSVFCQEQILVTE 206  
Db 1999 SCIQIIVGDVVSRGDINPTLKLTPLEOVLLNGLAINGFKLCKELIHHDVASGDDGLIN 2058  
QY 207 EVVVL 211  
Db 2059 SILIL 2063

Search completed: April 9, 2003, 14:33:39  
Job time : 32 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 9, 2003, 14:33:08 ; Search time 73 Seconds  
(without alignments)  
397.927 Million cell updates/sec

Title: US-10-034-500-2

Perfect score: 218  
Sequence: 1 AEVTASCTKRKVESNYLVLYD.....QEQLVTEVVLKGVNFAF 218

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 908470 seqs, 133250620 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

A\_Geneseq\_101002:.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3.7	85	23	ABG62118	Human prostete spe
2	3.2	74	21	ABP33399	Human ORF2372 prot
3	3.2	139	21	AAI2854	Arabidopsis thalia
4	3.2	141	21	AAI2853	Arabidopsis thalia
5	3.2	146	18	AAW20201	H. pylori secreted
6	3.2	148	21	AAI2852	Arabidopsis thalia
7	3.2	151	21	AAW25232	Eucalyptus grandis
8	3.2	157	21	AAW49083	Human transcriptio
9	3.2	159	22	ABG17907	Human transcriptio
10	3.2	168	21	AAI21701	Novel human diagno
					Arabidopsis thalia

11	3.2	170	21	AAI21700	Arabidopsis thalia
12	3.2	175	15	AAW62784	Borrelia M57 antiq
13	3.2	176	15	AAW62772	Borrelia B31 antiq
14	3.2	176	15	AAW62785	Borrelia W antigen
15	3.2	177	21	AAI2659	Arabidopsis thalia
16	3.2	191	15	AAW60898	Borrelia M57 antiq
17	3.2	192	15	AAW60899	Borrelia W antigen
18	3.2	192	15	AAW60886	Borrelia ip2 OspC
19	3.2	192	22	AAW62703	B burgdorferi ospc
20	3.2	192	22	AAW62721	B burgdorferi ospc
21	3.2	200	21	AAW28190	Arabidopsis thalia
22	3.2	200	22	AAW57004	Protonibacterium
23	3.2	205	23	ABP30264	Streptococcus poly
24	3.2	209	16	AAW75728	B. burgdorferi str
25	3.2	209	22	AAW62720	B burgdorferi ospc
26	3.2	210	16	AAW11935	Outer surface prot
27	3.2	210	16	AAW75727	B. burgdorferi str
28	3.2	228	21	AAW28189	Arabidopsis thalia
29	3.2	233	20	AAW07075	Renal cancer assoc
30	3.2	236	23	ABP26037	Streptococcus poly
31	3.2	238	20	AAW29487	Human lung tumour
32	3.2	238	21	AAW44412	Human lung tumour
33	3.2	238	22	AAW13753	Human lung tumour
34	3.2	243	20	AAW29489	Human lung tumour
35	3.2	243	20	AAW29491	Human lung tumour
36	3.2	243	21	AAW44414	Human lung tumour
37	3.2	243	21	AAW44417	Human lung tumour
38	3.2	243	22	AAW13755	Human lung tumour
39	3.2	243	22	AAW13758	Human lung tumour
40	3.2	244	20	AAW29495	Human lung tumour
41	3.2	244	20	AAW29492	Human lung tumour
42	3.2	244	21	AAW44418	Human lung tumour
43	3.2	244	21	AAW44421	Human lung tumour
44	3.2	244	22	AAW13759	Human lung tumour
45	3.2	245	22	AAW13762	Human lung tumour
46	3.2	245	20	AAW29490	Human lung tumour
47	3.2	245	21	AAW44415	Human lung tumour
48	3.2	245	22	AAW13756	Human lung tumour
49	3.2	257	22	AAW96279	Putative P. abyssi
50	3.2	258	22	ABG07381	Novel human diagno
51	3.2	278	19	AAW57572	Homo sapiens B268
52	3.2	311	22	AAW36280	Pseudomonas aerugi
53	3.2	312	21	AAW28188	Arabidopsis thalia
54	3.2	312	23	AAW91515	Outer surface prot
55	3.2	342	23	AAW91516	Outer surface prot
56	3.2	342	23	AAW91517	Outer surface prot
57	3.2	344	23	AAW91518	Outer surface prot
58	3.2	373	22	AAW62711	Borrelia sp chimera
59	3.2	374	22	AAW62710	Borrelia sp chimera
60	3.2	377	22	AAW62713	Borrelia sp chimera
61	3.2	378	22	AAW62712	Borrelia sp chimera
62	3.2	384	22	AAW62726	Borrelia sp chimera
63	3.2	387	22	AAW62729	Borrelia sp chimera
64	3.2	398	22	AAW62728	Borrelia sp chimera
65	3.2	400	22	AAW62739	Borrelia sp chimera
66	3.2	401	22	AAW62738	Borrelia sp chimera
67	3.2	402	21	AAW27429	Arabidopsis thalia
68	3.2	408	22	AAW62737	Borrelia sp chimera
69	3.2	409	23	ABW90217	Human polypeptide
70	3.2	429	19	AAW10949	H. pylori ORF 04cp
71	3.2	429	19	AAW46342	H. pylori HPS117 p
72	3.2	429	22	AAW46403	H. pylori HPC117 p
73	3.2	430	18	AAW20585	H. pylori secreted
74	3.2	440	23	AAW91519	OspC-B31/OspA-B31/
75	3.2	443	13	AAW27798	EBSP synthase wild
76	3.2	450	23	AAW91504	Outer surface prot
77	3.2	453	23	AAW91506	Outer surface prot
78	3.2	453	23	AAW91503	Outer surface prot
79	3.2	453	23	AAW91508	Outer surface prot
80	3.2	454	23	AAW91510	Outer surface prot
81	3.2	461	22	AAW41111	Protonibacterium
82	3.2	466	16	AAW75739	B31 outer surface
83	3.2	466	16	AAW75740	B31 outer surface

84	7	3.2	485	23	ABG64981	Human albumin fusi	157	6	2.8	63	22	AA673377	Human gene 32-enco
85	7	3.2	485	23	ABG64982	Human albumin fusi	158	6	2.8	63	22	AA673441	Human gene 32-enco
86	7	3.2	485	23	ABG64983	Human albumin fusi	159	6	2.8	63	22	AA673404	Human albumin fusi
87	7	3.2	485	23	AAE21466	Human gene 15 enco	160	6	2.8	63	23	ABG64305	Human albumin fusi
88	7	3.2	485	23	AAE21497	Human gene 15 enco	161	6	2.8	64	22	AAU47783	Propionibacterium
89	7	3.2	485	23	AAE21498	Human gene 15 enco	162	6	2.8	65	16	AAE82535	Hybrid IE beta cha
90	7	3.2	560	22	AAE67724	Borrelia sp chimex	163	6	2.8	65	22	AAE06578	Human foetal prote
91	7	3.2	587	16	AAE75746	B31 Osp-A/antigen	164	6	2.8	65	23	ABP09290	Human ORFX protein
92	7	3.2	658	22	ABG09545	Novel human diagno	165	6	2.8	68	22	AAU51370	Propionibacterium
93	7	3.2	749	23	ABB04872	LDL receptor bindi	166	6	2.8	70	22	ABG01420	Novel human diagno
94	7	3.2	852	22	AAE94405	Human protein sequ	167	6	2.8	70	22	ABE43646	Peptide #1152 enc
95	7	3.2	858	18	AAE23314	Human prostate pro	168	6	2.8	70	22	AAE64590	Human brain expres
96	7	3.2	858	20	AAE02436	N-terminal sequenc	169	6	2.8	70	22	AAE21313	Peptide #7747 enco
97	7	3.2	866	23	ABE49337	Listeria monocytog	170	6	2.8	70	23	ABG46403	Human peptide enco
98	7	3.2	986	19	AAE59050	H. contortus pHC-1	171	6	2.8	73	23	ABE32318	Human ORF1291 prot
99	7	3.2	986	23	AAO17501	H contortus HC110-	172	6	2.8	74	22	AAU41298	Propionibacterium
100	7	3.2	986	23	AAO17502	H contortus HC110-	173	6	2.8	75	22	AAU60214	Propionibacterium
101	7	3.2	986	23	AAO17513	H contortus HC110-	174	6	2.8	77	22	AAE85017	Human immune/haema
102	7	3.2	1005	21	AAE01833	Haemophilus influe	175	6	2.8	82	20	AAE99773	DEF chimeric molec
103	7	3.2	1011	21	AAE01832	Haemophilus influe	176	6	2.8	82	22	AAU49275	Propionibacterium
104	7	3.2	1077	23	ABE92833	Herbicidially activ	177	6	2.8	83	23	ABE34792	Human ORF1291 prot
105	7	3.2	1401	22	AAE94015	Human protein sequ	178	6	2.8	85	20	AAE35063	Strepococcus poly
106	7	3.2	1417	22	ABG09546	Novel human diagno	179	6	2.8	85	23	ABE29124	Human gene 12 enco
107	7	3.2	1678	22	ABE11539	Novel human diagno	180	6	2.8	87	22	AAE06116	Human gene 12 enco
108	7	3.2	2362	22	ABE58396	Drosophila melanog	181	6	2.8	87	22	AAE06117	C glutamincum prote
109	7	3.2	4345	22	ABE66417	UDP-D glucose:1lmo	182	6	2.8	87	22	AAE32990	Human secreted pro
110	6	2.8	14	22	AAE68121	Human peptide #477	183	6	2.8	87	23	ABG33938	Human secreted pro
111	6	2.8	14	22	AAE97202	Borrelia Ospc anti	184	6	2.8	87	23	ABG33939	Human secreted pro
112	6	2.8	15	22	AAE70361	Human DPP-3 pepti	185	6	2.8	87	23	ABE10366	Human ORFX protein
113	6	2.8	15	23	ABE61615	Human immune/haema	186	6	2.8	88	21	AAE77139	Human secreted pro
114	6	2.8	27	22	AAU89446	Insulin/insulin-11	187	6	2.8	88	21	AAE87140	Human secreted pro
115	6	2.8	28	22	ABE42683	Peptide #10189 enc	188	6	2.8	88	22	AAO05559	Human polypeptide
116	6	2.8	28	22	ABE26014	Protein #8013 enco	189	6	2.8	89	22	AAE94900	Human reproductive
117	6	2.8	28	22	AAE63573	Human brain expres	190	6	2.8	89	23	ABE07059	Human proteosome p
118	6	2.8	28	22	AAE76388	Human bone marrow	191	6	2.8	90	22	AAO12125	Human polypeptide
119	6	2.8	28	22	AAE64996	Peptide #10533 enc	192	6	2.8	91	22	AAE75579	Gene 28 human secr
120	6	2.8	28	22	ABG45674	Human peptide enco	193	6	2.8	93	21	AAE32637	Eucalyptus grandis
121	6	2.8	30	22	AAE84724	Human immune/haema	194	6	2.8	93	22	ABE31382	Peptide #4033 enco
122	6	2.8	34	22	ABE30719	Peptide #3370 enco	195	6	2.8	93	22	ABE36590	Peptide #4096 enco
123	6	2.8	34	22	ABE35895	Protein #3401 enco	196	6	2.8	93	22	ABE21930	Human brain expres
124	6	2.8	34	22	ABE21304	Human brain expres	197	6	2.8	93	22	AAE57353	Human bone marrow
125	6	2.8	34	22	AAE56700	Human bone marrow	198	6	2.8	93	22	AAE69750	Peptide #4003 enco
126	6	2.8	34	22	AAE69074	Peptide #3347 enco	199	6	2.8	93	22	AAE17569	Peptide #4129 enco
127	6	2.8	34	22	AAE16913	Peptide #3436 enco	200	6	2.8	93	22	AAE30092	Peptide #3919 enco
128	6	2.8	34	22	AAE29399	Peptide #3289 enco	201	6	2.8	93	22	AAE05237	Human peptide enco
129	6	2.8	34	22	AAE04607	Human peptide enco	202	6	2.8	93	23	ABE33976	Human ORFX protein
130	6	2.8	34	23	ABE38675	Human peptide enco	203	6	2.8	93	23	ABE04502	Human ORFX protein
131	6	2.8	37	16	AAE82534	IE beta chain frag	204	6	2.8	93	23	ABE06641	Human ORFX protein
132	6	2.8	42	22	AAE64697	Human secreted pro	205	6	2.8	93	23	ABE49894	Listeria monocytog
133	6	2.8	43	22	ABE27924	Human peptide #575	206	6	2.8	94	13	AAE21849	Sequence of Plasmid
134	6	2.8	43	22	ABE33098	Peptide #604 enco	207	6	2.8	94	21	AAE57412	Arabidopsis thalia
135	6	2.8	43	22	ABE18564	Protein #563 enco	208	6	2.8	94	22	AAU40923	Propionibacterium
136	6	2.8	43	22	AAE51895	Human brain expres	209	6	2.8	94	23	ABE35517	Human ORF4490 prot
137	6	2.8	43	22	AAE66283	Human bone marrow	210	6	2.8	95	23	ABE05525	Human ORFX protein
138	6	2.8	43	22	AAE14151	Peptide #585 enco	211	6	2.8	97	22	AAE60997	Propionibacterium
139	6	2.8	43	22	AAE26559	Peptide #596 enco	212	6	2.8	100	21	AAE35143	Arabidopsis thalia
140	6	2.8	43	22	AAE01891	Peptide #573 enco	213	6	2.8	100	22	AAE6961	Human prostate CDN
141	6	2.8	43	22	AAE76335	Human colon cancer	214	6	2.8	100	22	AAE01316	Pl1000C partial am1
142	6	2.8	43	23	ABE35930	Human peptide enco	215	6	2.8	100	23	ABE95421	Human pl1000C fragm
143	6	2.8	51	22	AAU46306	Propionibacterium	216	6	2.8	101	22	AAE23495	Novel human enzyme
144	6	2.8	52	22	AAU66151	Propionibacterium	217	6	2.8	101	22	AAO05134	Human polypeptide
145	6	2.8	53	23	ABE07055	Human ORFX protein	218	6	2.8	103	21	AAE52091	Gene 39 human secr
146	6	2.8	55	21	AAE33258	Human radiata tran	219	6	2.8	104	19	AAE79457	Staphylococcus aur
147	6	2.8	58	19	AAE75165	Human secreted pro	220	6	2.8	105	14	AAE30619	Polypeptide coded
148	6	2.8	58	21	AAE32894	Human radiata tran	221	6	2.8	106	21	AAE57622	Arabidopsis thalia
149	6	2.8	59	19	AAE75109	Human secreted pro	222	6	2.8	106	21	AAE59992	Arabidopsis thalia
150	6	2.8	59	21	AAE09734	Arabidopsis thalia	223	6	2.8	106	22	AAU20292	Human novel endoc
151	6	2.8	59	21	AAE55025	Arabidopsis thalia	224	6	2.8	107	23	ABE76264	Rabbit bone marrow
152	6	2.8	61	22	ABE03453	Novel human diagno	225	6	2.8	108	21	AAE33641	Arabidopsis thalia
153	6	2.8	62	22	ABE42312	Peptide #9718 enco	226	6	2.8	109	21	AAE18130	Human brain expres
154	6	2.8	62	22	AAE63097	Human brain expres	227	6	2.8	109	22	AAE08811	Human haematologic
155	6	2.8	62	22	AAE75908	Human bone marrow	228	6	2.8	111	20	AAU4769	Mycobacterium spec
156	6	2.8	62	22	AAE36019	Peptide #10056 enc	229	6	2.8	111	21	AAE52505	Helicobacter pylor

230	6	2.8	117	21	AGI5142	Arabidopsis thalia	303	6	2.8	186	21	AG33664	Arabidopsis thalia
231	6	2.8	118	21	AG09732	Arabidopsis thalia	304	6	2.8	187	21	AG15918	Arabidopsis thalia
232	6	2.8	118	21	AG50523	Arabidopsis thalia	305	6	2.8	187	21	AG34430	Arabidopsis thalia
233	6	2.8	119	22	AAE07116	Human gene 10 enco	306	6	2.8	187	21	AG38061	Arabidopsis thalia
234	6	2.8	120	23	ABP07019	Human ORF protein	307	6	2.8	187	21	AG33002	Arabidopsis thalia
235	6	2.8	121	21	AAE32750	Eucalyptus grandis	308	6	2.8	188	21	AAE5164	Human secreted pro
236	6	2.8	121	22	ABE17222	Human nervous syst	309	6	2.8	188	21	AG39621	Arabidopsis thalia
237	6	2.8	122	22	AAU03701	Group B Streptococ	310	6	2.8	188	22	ABE52878	Escherichia coli P
238	6	2.8	122	23	AAE24065	Human prostate spe	311	6	2.8	189	21	AG06441	Arabidopsis thalia
239	6	2.8	125	21	AG33640	Arabidopsis thalia	312	6	2.8	191	15	AAE60897	Borrelia VS461 ant
240	6	2.8	125	23	AAE51937	Human TGFbeta prot	313	6	2.8	192	21	AAE94941	Human secreted pro
241	6	2.8	126	21	AAE32660	Eucalyptus grandis	314	6	2.8	193	22	AAE65763	S. pombe protein K
242	6	2.8	126	22	AAU57698	Propionibacterium	315	6	2.8	194	15	AAE60892	Borrelia ORF anti
243	6	2.8	127	22	AAE78729	Human protein SEQ	316	6	2.8	195	22	AAU48915	Propionibacterium
244	6	2.8	131	22	AAE39931	Human polypeptide	317	6	2.8	197	20	AAE96222	Smad1 protein C-te
245	6	2.8	131	22	AAE60494	Human cell cycle a	318	6	2.8	197	20	AAE96226	Mad protein C-term
246	6	2.8	132	21	AAE18010	Pinus radiata pine	319	6	2.8	197	21	AAE24767	Plant SDF encoded
247	6	2.8	132	22	ABE11412	Human secreted pro	320	6	2.8	198	20	AAE96225	Smad5 protein C-te
248	6	2.8	132	22	AAE31378	C glutamicum prote	321	6	2.8	199	20	AAE96244	Smad1 protein frag
249	6	2.8	134	22	AAE55578	Propionibacterium	322	6	2.8	199	20	AAE09552	Arabidopsis thalia
250	6	2.8	135	22	AAE25882	Human protein sequ	323	6	2.8	202	21	AG330163	Arabidopsis thalia
251	6	2.8	136	23	ABP11174	Human ORF protein	324	6	2.8	202	21	AG34429	Arabidopsis thalia
252	6	2.8	137	14	AAE32996	Partial P. falcipar	325	6	2.8	202	21	AAE38060	Arabidopsis thalia
253	6	2.8	137	20	AAE09038	BSP-GST 3' deletio	326	6	2.8	202	22	AAE09290	Fibroblast growth
254	6	2.8	139	22	ABE10252	Propionibacterium	327	6	2.8	202	22	AAE85833	Murine fibroblast
255	6	2.8	139	22	AAE65712	Human cDNA SEQ ID	328	6	2.8	202	22	AAE45801	Rat KFGF protein.
256	6	2.8	140	21	ABE19777	Arabidopsis thalia	329	6	2.8	204	21	AG33563	Arabidopsis thalia
257	6	2.8	142	22	AAE62254	Propionibacterium	330	6	2.8	207	16	AAE77169	Arabidopsis thalia
258	6	2.8	143	21	AAE54955	Arabidopsis thalia	331	6	2.8	207	20	AAE21990	Human complement f
259	6	2.8	147	22	AAE87796	Human immune/haema	332	6	2.8	207	21	AG39620	Arabidopsis thalia
260	6	2.8	147	22	AAE00034	Human polypeptide	333	6	2.8	208	19	AAE98565	H. pylori GHP 179
261	6	2.8	150	22	ABE69536	Drosophila melanog	334	6	2.8	209	22	AAE69940	Human transmembran
262	6	2.8	150	22	AAU37842	Streptococcus pneu	335	6	2.8	209	22	AAE01295	Human transmembran
263	6	2.8	151	23	AAE01750	Human nucleoporin	336	6	2.8	209	23	AAE95400	Human transmembran
264	6	2.8	154	22	AAE88493	Haemophilus influe	337	6	2.8	212	14	AAE30909	B. burgdorferi pc
265	6	2.8	154	23	AAU91424	Haemophilus influe	338	6	2.8	212	21	AAE06440	Arabidopsis thalia
266	6	2.8	155	22	ABE12375	Human bone marrow	339	6	2.8	214	23	ABP05353	Human ORF protein
267	6	2.8	157	21	AAE19776	Arabidopsis thalia	340	6	2.8	216	22	AAE93722	Human protein sequ
268	6	2.8	159	14	AAE32585	P. falciparum EB200	341	6	2.8	220	22	ABE08235	Novel human diagno
269	6	2.8	159	14	AAE32993	P. falciparum EB200	342	6	2.8	220	22	AAE56605	Human brain expres
270	6	2.8	160	21	AAE06442	Arabidopsis thalia	343	6	2.8	221	22	ABE06533	Novel human diagno
271	6	2.8	161	21	AAE57621	Arabidopsis thalia	344	6	2.8	222	19	AAE98524	H. pylori GHP 137
272	6	2.8	161	21	AAE59991	Arabidopsis thalia	345	6	2.8	223	21	AAE09551	Arabidopsis thalia
273	6	2.8	161	22	AAE41717	Human polypeptide	346	6	2.8	223	22	AAE48369	Human SEC2 protein
274	6	2.8	162	23	ABP30380	Streptococcus poly	347	6	2.8	223	23	ABE54965	Lactococcus lactis
275	6	2.8	163	21	AAE59980	Arabidopsis thalia	348	6	2.8	224	19	AAE51121	Mouse endymnin-1i
276	6	2.8	163	22	AAU18031	Human immunoglobul	349	6	2.8	224	21	AAE15621	Arabidopsis thalia
277	6	2.8	164	21	AAE51745	Human secreted pro	350	6	2.8	224	21	AAE94655	Murine homology to
278	6	2.8	165	20	AAE04770	Mycobacterium spec	351	6	2.8	225	21	AAE32401	Arabidopsis thalia
279	6	2.8	166	22	ABE07311	Novel human diagno	352	6	2.8	225	21	AAE47382	Arabidopsis thalia
280	6	2.8	169	21	AAE41371	Human ORF ORF135	353	6	2.8	226	23	ABP38100	Staphylococcus epi
281	6	2.8	169	23	ABP32470	Human ORF143 prot	354	6	2.8	227	21	AAE10079	Arabidopsis thalia
282	6	2.8	171	21	AAE57620	Arabidopsis thalia	355	6	2.8	227	23	ABE92506	Herbicidally activ
283	6	2.8	174	23	ABE28355	Streptococcus poly	356	6	2.8	229	13	AAE21613	Sequence of Plasm
284	6	2.8	175	15	AAE2783	Borrelia VS461 ant	357	6	2.8	230	22	AAE6724	Corynebacterium gl
285	6	2.8	175	16	AAE73018	Pig interleukin-10	358	6	2.8	231	21	AAE28205	Arabidopsis thalia
286	6	2.8	175	16	AAE71471	Porcine interleuki	359	6	2.8	231	23	ABE48736	Listeria monocytog
287	6	2.8	175	21	AAE34431	Arabidopsis thalia	360	6	2.8	232	23	ABE78815	Human NOV7 protein
288	6	2.8	175	21	AAE38062	Arabidopsis thalia	361	6	2.8	232	22	AAE91894	C glutamicum prote
289	6	2.8	177	22	ABE17711	Human nervous syst	362	6	2.8	233	23	ABE30978	Extracellular and
290	6	2.8	177	22	AAE48073	Human extracellular	363	6	2.8	234	23	ABE47489	Listeria monocytog
291	6	2.8	178	15	AAE2778	Borrelia ORF anti	364	6	2.8	236	22	AAE58585	Propionibacterium
292	6	2.8	178	21	AAE52429	Staphylococcus aur	365	6	2.8	238	21	AAE21516	Arabidopsis thalia
293	6	2.8	178	23	ABP10979	Predicted extracel	366	6	2.8	238	21	AAE28204	Arabidopsis thalia
294	6	2.8	179	21	AAE24768	Plant SDF encoded	367	6	2.8	238	22	ABE11062	Human DNA binding
295	6	2.8	180	21	AAE54744	Arabidopsis thalia	368	6	2.8	240	22	AAE14141	Human novel protei
296	6	2.8	180	22	ABE16247	Novel human diagno	369	6	2.8	240	22	AAE94472	Human protein sequ
297	6	2.8	181	21	AAE41845	Human ORF ORF1609	370	6	2.8	242	20	AAE44179	N.meningitidis nla
298	6	2.8	181	21	AAE67075	Human secreted pro	371	6	2.8	243	21	AAE28203	Arabidopsis thalia
299	6	2.8	181	22	AAE60652	Human gene 12 enco	372	6	2.8	244	22	AAE80045	Corynebacterium gl
300	6	2.8	181	23	ABE33874	Human secreted pro	373	6	2.8	246	21	AAE75088	Neisseria gonorrhoe
301	6	2.8	182	19	AAE56799	S. thermophilus fl	374	6	2.8	246	23	ABE09647	Amino acid sequenc
302	6	2.8	182	22	AAE52222	Propionibacterium	375	6	2.8	247	21	AAE32632	Eucalyptus grandis

376	6	2.8	247	21	AAV75089	Neisseria meningit	449	6	2.8	296	22	AAU00498	Human TANGO 315 fo
377	6	2.8	247	21	AAV75090	Neisseria meningit	450	6	2.8	296	23	ABP26184	Streptococcus poly
378	6	2.8	247	22	AAE12020	Medicago sativa ca	451	6	2.8	297	22	ABBE9196	Drosophila melanog
379	6	2.8	248	19	AAAG4265	Xylanase activity	452	6	2.8	297	23	ABBE6140	Human NS protein s
380	6	2.8	250	21	AAAG0470	Arabidopsis thalia	453	6	2.8	298	21	AAV97220	Stillic acid bindin
381	6	2.8	250	21	AAAG34920	Arabidopsis thalia	454	6	2.8	298	21	AAAG3558	Arabidopsis thalia
382	6	2.8	250	21	AAU78810	Autoinducer Inacti	455	6	2.8	298	21	AAAG3570	Arabidopsis thalia
383	6	2.8	254	16	AAAR82532	Hybrid 1E beta cha	456	6	2.8	299	21	AAAB5317	Musca mulatta rha
384	6	2.8	255	21	AAAG21515	Arabidopsis thalia	457	6	2.8	300	22	ABG11897	Novel human diagno
385	6	2.8	256	10	AAAP90531	Hypodermis B of Ar	458	6	2.8	302	22	ABBE7860	Drosophila melanog
386	6	2.8	257	21	AAAG21514	Arabidopsis thalia	459	6	2.8	303	21	AAAG6786	Arabidopsis thalia
387	6	2.8	258	23	ABBE98214	Chlamydia polytyp	460	6	2.8	303	23	ABBA9742	Listeria monocytog
388	6	2.8	259	21	AAAG30162	Arabidopsis thalia	461	6	2.8	304	21	AAAG6086	Arabidopsis thalia
389	6	2.8	261	22	AAAB39364	Human polypeptide	462	6	2.8	304	23	AAU93017	Arabidopsis transc
390	6	2.8	261	22	AAAB94185	Human protein sequ	463	6	2.8	309	22	ABG01154	Novel human diagno
391	6	2.8	261	23	ABBT89888	Lolium perenne	464	6	2.8	312	20	AAV31794	Porcine circovirus
392	6	2.8	261	23	ABBT89889	Lolium perenne CCo	465	6	2.8	312	20	AAV33624	Protein encoded by
393	6	2.8	262	21	AAAG09550	Arabidopsis thalia	466	6	2.8	312	21	AAAG0549	Arabidopsis thalia
394	6	2.8	262	23	ABBT07984	C. glutamicum poly	467	6	2.8	312	21	AAAG34919	Arabidopsis thalia
395	6	2.8	263	22	AAAG89806	C. glutamicum prote	468	6	2.8	312	22	AAU35746	Helicobacter pylor
396	6	2.8	264	21	AAV68288	Class II beta chat	469	6	2.8	312	22	AAU35918	Helicobacter pylor
397	6	2.8	264	21	AAV52942	Class II beta chat	470	6	2.8	313	21	AAAG53569	Arabidopsis thalia
398	6	2.8	264	22	ABG00508	Novel human diagno	471	6	2.8	313	23	ABBA9438	Listeria monocytog
399	6	2.8	264	22	ABG21392	Novel human diagno	472	6	2.8	315	23	AAAT9892	Hepatitis C virus
400	6	2.8	264	22	AAAB58703	Class II beta chat	473	6	2.8	316	21	AAAG50782	Arabidopsis thalia
401	6	2.8	265	21	ABP30234	Streptococcus poly	474	6	2.8	317	22	ABG50782	Psychrobacter immo
402	6	2.8	266	21	AAAB32853	Ecicalyptus grandis	475	6	2.8	317	22	ABBB09184	Psychrobacter immo
403	6	2.8	266	22	ABG02477	Novel human diagno	476	6	2.8	317	22	AAAG72998	Olfactory receptor
404	6	2.8	267	18	AAAM23303	Rat CRTI. Rattus	477	6	2.8	318	23	ABP11025	Amino acid sequenc
405	6	2.8	267	23	ABBT89858	Lolium perenne lpc	478	6	2.8	320	19	AAAM7422	Bacillus strearoth
406	6	2.8	268	21	AAV44848	Triticum aestivum	479	6	2.8	320	22	ABBT1307	Drosophila melanog
407	6	2.8	268	22	ABBE62086	Drosophila melanog	480	6	2.8	323	17	AAAR2060	Hepaprenyl diaphos
408	6	2.8	269	22	AAAB76538	Corynebacterium gl	481	6	2.8	324	21	AAAG33557	Arbiodopsis thalia
409	6	2.8	270	21	AAAG07724	Arabidopsis thalia	482	6	2.8	325	23	ABBA9099	Listeria monocytog
410	6	2.8	270	21	AAAG46796	Glycerol-3-phospha	483	6	2.8	331	20	AAAG67889	Human secreted pro
411	6	2.8	271	19	AAAM30683	Klebsiella pneumon	484	6	2.8	333	21	AAAG05468	Arabidopsis thalia
412	6	2.8	271	19	AAAG60263	Mitochondrial glyco	485	6	2.8	333	21	AAAG20277	Arabidopsis thalia
413	6	2.8	271	19	AAAM57327	Cytosolic glycerol	486	6	2.8	333	21	AAAG34918	Arabidopsis thalia
414	6	2.8	271	20	AAV26169	Arabidopsis thalia	487	6	2.8	333	21	AAAG34039	Arabidopsis thalia
415	6	2.8	271	21	AAAG06087	Arabidopsis thalia	488	6	2.8	336	20	AAV37444	Protein involved 1
416	6	2.8	272	21	AAAG53571	Arabidopsis thalia	489	6	2.8	338	21	AAAG47141	Arabidopsis thalia
417	6	2.8	272	22	ABBE63475	Drosophila melanog	490	6	2.8	338	22	AAUS3725	Propionibacterium
418	6	2.8	274	22	ABG15472	Novel human diagno	491	6	2.8	339	21	AAAG24300	Arabidopsis thalia
419	6	2.8	274	22	AAAB96687	Putative P. abyssi	492	6	2.8	339	22	AAU32845	Novel human secret
420	6	2.8	275	22	ABBO9061	Thermus caldophilu	493	6	2.8	342	22	AAAB96175	Putative P. abyssi
421	6	2.8	275	22	ABG03454	Novel human diagno	494	6	2.8	344	12	AAAR10864	Lipase modulatiog
422	6	2.8	276	22	AAUS2671	Propionibacterium	495	6	2.8	344	14	AAAR39397	Pseudomonas cepact
423	6	2.8	276	22	AAAG91845	C. glutamicum prote	496	6	2.8	344	16	AAAR72246	P. cepactia lipase
424	6	2.8	276	22	AAAB61617	Human protein HP10	497	6	2.8	345	22	AAU33858	Staphylococcus aur
425	6	2.8	277	20	AAV34592	Chlamydia pneumoni	498	6	2.8	347	22	AAAG2805	C. glutamicum prote
426	6	2.8	277	22	AAU00499	Human TANGO 315 fo	499	6	2.8	347	19	AAAG60619	Human liver activi
427	6	2.8	278	22	AAAG8326	S. epidermidis ope	500	6	2.8	352	22	ABBS9023	Drosophila melanog
428	6	2.8	280	21	AAAG46787	Arabidopsis thalia	501	6	2.8	353	22	AAAG3289	Human polypeptide
429	6	2.8	280	21	AAAG47143	Arabidopsis thalia	502	6	2.8	354	21	AAAB42518	Human ORFX ORF282
430	6	2.8	282	18	AAAM1261	Human Bmp-15 prote	503	6	2.8	354	21	AAAG07722	Arabidopsis thalia
431	6	2.8	282	21	AAAG30161	Arabidopsis thalia	504	6	2.8	354	21	AAAG6794	Arabidopsis thalia
432	6	2.8	283	22	AAAB96242	Putative P. abyssi	505	6	2.8	356	22	AAU00858	S. aureus D-alanin
433	6	2.8	283	23	AAAM52833	Physcomitrella pat	506	6	2.8	357	18	AAAM4337	Tomato yellow leaf
434	6	2.8	286	21	AAV97221	Stillic acid bindin	507	6	2.8	357	18	AAAM4329	Tomato yellow leaf
435	6	2.8	286	21	AAAG14019	Arabidopsis thalia	508	6	2.8	357	18	AAAM4330	Tomato yellow leaf
436	6	2.8	286	21	AAAG45986	Arabidopsis thalia	509	6	2.8	357	18	AAAM4331	Tomato yellow leaf
437	6	2.8	286	22	AAAG92094	C. glutamicum prote	510	6	2.8	358	21	AAAG24299	Arabidopsis thalia
438	6	2.8	286	22	AAAG82614	S. epidermidis ope	511	6	2.8	358	22	AAU00784	Human apocpobis pr
439	6	2.8	287	21	AAAG47142	Arabidopsis thalia	512	6	2.8	360	20	AAV08451	Maize Ubi1 homolo
440	6	2.8	287	22	AAAM41150	Human polypeptide	513	6	2.8	360	21	AAV77470	Human deubiquitina
441	6	2.8	289	22	AAAB49668	Psychrobacter immo	514	6	2.8	361	8	AAV70562	Product of ORF 4 f
442	6	2.8	291	21	ABBS3757	Lactococcus lactis	515	6	2.8	363	21	AAAG33536	Arabidopsis thalia
443	6	2.8	293	21	AAAG07723	Arabidopsis thalia	516	6	2.8	363	23	ABBS90963	Herbicially activ
444	6	2.8	293	21	AAAG46795	Arabidopsis thalia	517	6	2.8	364	21	AAAG07248	Arabidopsis thalia
445	6	2.8	294	21	AAAB07857	Amino acid sequenc	518	6	2.8	364	21	AAAG46785	Arabidopsis thalia
446	6	2.8	294	23	ABBS07883	Arabidopsis thalia	519	6	2.8	366	23	ABP27228	Streptococcus poly
447	6	2.8	294	23	ABBS47859	Lactococcus lactis	520	6	2.8	367	22	AAE12631	Human gene 1 encod
448	6	2.8	295	21	AAV97219	Stillic acid bindin	521	6	2.8	369	21	AAAG24298	Arabidopsis thalia

522	6	2.8	370	22	ABR30007	Peptide #2658 enco	595	6	2.8	442	22	ABR66680	Drosophila melanog
523	6	2.8	370	22	ABR20615	Protein #2614 enco	596	6	2.8	445	22	AAU20465	Human secreted pro
524	6	2.8	370	22	AAM56010	Human brain expres	597	6	2.8	446	20	AAU29179	Amino acid sequenc
525	6	2.8	370	22	AAM68380	Human bone marrow	598	6	2.8	446	20	AAU18026	Murine DP-3alpha i
526	6	2.8	370	22	AAM16198	Peptide #2632 enco	599	6	2.8	447	14	AAU32656	C. murinae ATCC
527	6	2.8	370	22	AAM03928	Peptide #2610 enco	600	6	2.8	447	22	AAU93228	C. glutamicum prote
528	6	2.8	370	23	ABG37945	Human peptide enco	601	6	2.8	447	22	ABR83187	Brevibacterium lac
529	6	2.8	370	23	ABR48054	Listeria monocytog	602	6	2.8	447	22	ABR79680	Corynebacterium gl
530	6	2.8	372	22	ABR58642	Drosophila melanog	603	6	2.8	449	23	ABP41485	Human ovarian anti
531	6	2.8	373	22	ABR60156	Drosophila melanog	604	6	2.8	450	21	AAU21555	Arabidopsis thalia
532	6	2.8	374	22	AAU37880	Streptococcus pneu	605	6	2.8	452	22	ABR96096	Purative P. abyssi
533	6	2.8	374	22	AAU38057	Streptococcus pneu	606	6	2.8	452	23	ABR83254	Yeast prephenate d
534	6	2.8	374	22	AAU72131	Safflower plactidi	607	6	2.8	454	22	AAU39074	Human polyphenide
535	6	2.8	374	22	AAU72132	Safflower plactidi	608	6	2.8	454	22	AAU93774	Human protein sequ
536	6	2.8	375	22	ABR83046	Human WPI1 HECT E3	609	6	2.8	454	22	ABR68521	Human GRP-binding
537	6	2.8	377	21	AAU07247	Arabidopsis thalia	610	6	2.8	456	21	AAU50525	Arabidopsis thalia
538	6	2.8	378	21	AAU40108	Arabidopsis thalia	611	6	2.8	460	21	AAU43157	Human ORF1 ORF2921
539	6	2.8	378	21	AAU46985	Arabidopsis thalia	612	6	2.8	460	21	AAU35697	Arabidopsis thalia
540	6	2.8	378	21	ABR93965	Herbicideally activ	613	6	2.8	460	21	AAU37220	Xenopus alpha-sig
541	6	2.8	379	23	ABP27052	Streptococcus poly	614	6	2.8	464	18	AAU18094	Human alpha-sig
542	6	2.8	380	21	AAU50781	Human WPI1 HECT E3	615	6	2.8	465	18	AAU18098	Human alpha-sig
543	6	2.8	380	22	ABR83047	Human WPI1 HECT E3	616	6	2.8	465	21	AAU17563	Arabidopsis thalia
544	6	2.8	384	23	ABR91551	Herbicideally activ	617	6	2.8	465	21	AAU81444	Human Smad5. Homo
545	6	2.8	387	22	AAU61205	Mycobacterium tube	618	6	2.8	465	21	AAU76806	Human Smad1. protei
546	6	2.8	387	22	ABP39453	Staphylococcus epi	619	6	2.8	465	22	AAU76806	Human Smad1. protei
547	6	2.8	392	21	AAU92031	Human bone morphog	620	6	2.8	468	21	AAU21705	Novel human neopla
548	6	2.8	393	22	AAU36802	Staphylococcus aur	621	6	2.8	470	21	AAU30955	Arabidopsis thalia
549	6	2.8	398	21	AAU07246	Arabidopsis thalia	622	6	2.8	474	20	AAU47134	Amino acid sequenc
550	6	2.8	398	22	AAU06622	Human secreted pro	623	6	2.8	474	20	AAU00180	Enterococcus faeca
551	6	2.8	406	21	AAU94209	Human TRAF four as	624	6	2.8	474	23	ABP43399	E. faecalis EPO90 p
552	6	2.8	407	21	AAU06085	Arabidopsis thalia	625	6	2.8	475	22	AAU98410	Escherichia coli p
553	6	2.8	411	23	ABP27251	Streptococcus poly	626	6	2.8	477	21	AAU47133	Arabidopsis thalia
554	6	2.8	412	22	AAU86090	Putative P. abyssi	627	6	2.8	479	21	AAU50524	Arabidopsis thalia
555	6	2.8	414	21	AAU858368	Lung cancer associ	628	6	2.8	482	21	AAU47132	Arabidopsis thalia
556	6	2.8	414	21	AAU807856	Amino acid sequenc	629	6	2.8	484	22	AAU92676	C. glutamicum prote
557	6	2.8	414	23	ABR41711	Human ovarian anti	630	6	2.8	484	22	AAU93036	Human protein sequ
558	6	2.8	415	21	AAU21556	Arabidopsis thalia	631	6	2.8	484	22	AAU79356	Corynebacterium gl
559	6	2.8	415	22	AAU56212	Propionibacterium	632	6	2.8	487	22	AAU72130	Safflower plactidi
560	6	2.8	417	19	AAU54355	47 KD heat shock p	633	6	2.8	488	21	AAU17562	Arabidopsis thalia
561	6	2.8	417	21	AAU81908	Human Hsp47 protei	634	6	2.8	489	21	AAU20803	Arabidopsis thalia
562	6	2.8	418	23	ABR91469	Herbicideally activ	635	6	2.8	489	21	AAU51384	Arabidopsis thalia
563	6	2.8	418	23	AAU75577	Human heat shock p	636	6	2.8	489	22	ABR62027	Drosophila melanog
564	6	2.8	422	17	AAU89906	Human kynurenine a	637	6	2.8	490	23	ABR60306	Lymphoma associate
565	6	2.8	422	21	AAU96722	E. gracilis fatty	638	6	2.8	491	22	ABR52622	Escherichia coli p
566	6	2.8	424	22	ABR57933	Drosophila melanog	639	6	2.8	492	21	AAU36901	Human TMPRSS2. prot
567	6	2.8	424	23	ABP40105	Staphylococcus epi	640	6	2.8	492	21	AAU21554	Arabidopsis thalia
568	6	2.8	426	21	AAU80344	Candida albicans e	641	6	2.8	492	21	AAU35696	Arabidopsis thalia
569	6	2.8	426	22	ABR79858	Corynebacterium gl	642	6	2.8	492	21	AAU37219	Arabidopsis thalia
570	6	2.8	427	21	AAU07585	Guar phosphomanno	643	6	2.8	492	21	AAU92050	Arabidopsis thalia
571	6	2.8	428	18	AAU26603	Enterococcus faeca	644	6	2.8	492	21	AAU57280	Ovrl15 homolog pro
572	6	2.8	431	20	AAU00181	E. faecalis EPO90 a	645	6	2.8	492	21	AAU77726	Human tumor suppr
573	6	2.8	431	23	ABP43400	Herbicideally activ	646	6	2.8	492	21	AAU44406	Human 20P1F12-GTC2
574	6	2.8	431	23	ABR90809	Brevibacterium lac	647	6	2.8	492	22	ABR63749	Drosophila melanog
575	6	2.8	432	19	AAU69554	C. glutamicum prote	648	6	2.8	492	22	AAU69939	Human transmembran
576	6	2.8	432	22	AAU30017	Arabidopsis thalia	649	6	2.8	492	22	AAU69930	Human serine prote
577	6	2.8	434	21	AAU30957	Human cancer assoc	650	6	2.8	492	22	AAU69943	Human transmembran
578	6	2.8	436	21	AAU43483	Glycine max 3-dehy	651	6	2.8	492	22	AAU013294	P1000C amino acid
579	6	2.8	436	21	AAU44846	Human ovarian carc	652	6	2.8	492	22	AAU01315	Prostate cancer-as
580	6	2.8	437	21	AAU44847	Zea mays 3-dehydro	653	6	2.8	492	23	ABR61885	Human transmembran
581	6	2.8	438	21	AAU12555	Escherichia coli p	654	6	2.8	492	23	ABR95420	Human P1000C SEQ I
582	6	2.8	438	22	ABR52864	Human gene 1. encod	655	6	2.8	492	23	AAU18096	Human 20P1F12-TMPC
583	6	2.8	438	22	AAU12623	0772P clone 21008.	656	6	2.8	492	23	AAU18097	Human 20P1F12-TMPC
584	6	2.8	438	22	ABR30989	Partial protein ee	657	6	2.8	492	23	AAU18099	Human 20P1F12-TMPC
585	6	2.8	438	23	ABR30973	Anticentigen diagno	658	6	2.8	492	23	AAU18099	Human 20P1F12-TMPC
586	6	2.8	439	21	AAU95851	Arabidopsis thalia	659	6	2.8	492	23	AAU18100	Human 20P1F12-TMPC
587	6	2.8	439	21	AAU30956	Novel human diagno	660	6	2.8	492	23	AAU18100	Human 20P1F12-TMPC
588	6	2.8	441	22	ABG11753	Arabidopsis thalia	661	6	2.8	494	18	AAU30711	Human ubiqutin-SP
589	6	2.8	442	21	AAU40107	Arabidopsis thalia	662	6	2.8	496	23	AAU47798	LFKS3 condensin
590	6	2.8	442	21	AAU40804	Arabidopsis thalia	663	6	2.8	497	20	AAU88329	Arabidopsis very 1
591	6	2.8	442	21	AAU32486	Arabidopsis thalia	664	6	2.8	497	21	AAU35695	Arabidopsis thalia
592	6	2.8	442	21	AAU46984	Arabidopsis thalia	665	6	2.8	497	21	AAU37218	Arabidopsis thalia
593	6	2.8	442	21	AAU46984	Arabidopsis thalia	666	6	2.8	497	21	AAU37218	Arabidopsis thalia
594	6	2.8	442	21	AAU51385	Arabidopsis thalia	667	6	2.8	500	20	AAU93432	A. thaliana Elv6 pr

668	6	2.8	505	16	AAR77172	Condensing enzyme	741	6	2.8	638	22	ABR70740	Drosophila melanog
669	6	2.8	506	17	AAR95594	Arbidopsis fatty	742	6	2.8	645	19	AAW42084	Amino acid sequenc
670	6	2.8	506	23	ABR48158	Listeria monocytog	743	6	2.8	645	22	AAE02547	A. thaliana transc
671	6	2.8	506	23	AAE17608	Arbidopsis thalia	744	6	2.8	645	22	AAE01903	Arbidopsis thalia
672	6	2.8	506	23	AAE17611	A. thaliana PAEI-B	745	6	2.8	645	22	AAE01926	Arbidopsis thalia
673	6	2.8	506	23	AAE17613	A. thaliana PAEI-B	746	6	2.8	645	23	AAU33022	Arbidopsis transc
674	6	2.8	506	23	AAE17614	A. thaliana PAEI-B	747	6	2.8	654	21	AAV53013	Human secreted pro
675	6	2.8	506	23	AAE17615	A. thaliana PAEI-B	748	6	2.8	654	22	AAE95459	Human protein sequ
676	6	2.8	506	23	AAE17617	A. thaliana PAEI-B	749	6	2.8	661	19	ABBS9862	Drosophila melanog
677	6	2.8	506	23	AAE17618	A. thaliana PAEI-B	750	6	2.8	662	19	AAW98677	H. pylori GPHO 564
678	6	2.8	506	23	AAE17619	A. thaliana PAEI-B	751	6	2.8	662	22	AAU55792	Helicobacter pylor
679	6	2.8	506	23	AAE17620	A. thaliana PAEI-B	752	6	2.8	666	22	ABBS5593	Drosophila melanog
680	6	2.8	506	23	AAE17622	A. thaliana PAEI-B	753	6	2.8	667	23	ABBS5098	Lactococcus lactis
681	6	2.8	506	23	AAE17625	Arbidopsis thalia	754	6	2.8	675	23	ABBS0767	Human Tumour Endot
682	6	2.8	506	23	AAE17626	A. thaliana PAEI-B	755	6	2.8	677	23	AAU36507	Pseudomonas aerugi
683	6	2.8	506	23	AAE17627	A. thaliana PAEI-B	756	6	2.8	680	20	AAV13451	Amino acid sequenc
684	6	2.8	506	23	AAE17848	Alternative versio	757	6	2.8	683	18	AAW6794	Novel human protei
685	6	2.8	506	23	AAE17849	Alternative versio	758	6	2.8	683	22	AAE05494	Human ubiquitin pr
686	6	2.8	506	23	AAE17850	Alternative versio	759	6	2.8	684	22	AAW78967	Drosophila melanog
687	6	2.8	507	22	ABG25482	Novel human diagno	760	6	2.8	687	21	AAV78981	Silkworm Bm white
688	6	2.8	507	22	AAE60098	Human transport pr	761	6	2.8	687	22	ABBS9384	Drosophila melanog
689	6	2.8	507	22	AAE16940	UDP-glucuronosyl c	762	6	2.8	689	22	AAW41382	Human polypeptide
690	6	2.8	510	22	AAE93044	C. glutamicum prote	763	6	2.8	689	22	AAW41383	Human polypeptide
691	6	2.8	510	22	AAE84890	Zinc finger protei	764	6	2.8	691	23	ABG61612	Human DRP-3 splic
692	6	2.8	510	22	AAE14681	Human transcripctio	765	6	2.8	693	23	ABB77503	Corynebacterium He
693	6	2.8	511	22	AAE68119	Citrus unshiu UDP-	766	6	2.8	706	23	ABG61611	Human DRP-3 splic
694	6	2.8	511	22	AAE68123	Citrus unshiu UDP-	767	6	2.8	708	20	AAV13452	Amino acid sequenc
695	6	2.8	512	18	AAW35704	Snaptagon flavono	768	6	2.8	715	22	ABG27848	Novel human diagno
696	6	2.8	512	21	AAW50523	Arbidopsis thalia	769	6	2.8	724	23	ABP26011	Streptococcus poly
697	6	2.8	512	21	AAU35639	Haemophilus influe	770	6	2.8	733	22	ABBS0048	Human clone 811a
698	6	2.8	513	22	ABR91770	Herbicidally activ	771	6	2.8	739	22	ABG16477	Novel human diagno
699	6	2.8	513	22	AAU27726	Human full-length	772	6	2.8	739	22	AAE68173	Atrophin-1 interac
700	6	2.8	513	22	ABR97230	Novel human protei	773	6	2.8	739	22	ABP30271	Streptococcus poly
701	6	2.8	516	20	AAV07735	Human breast-speci	774	6	2.8	747	23	ABP25975	Streptococcus poly
702	6	2.8	517	22	AAE65240	Cell division cont	775	6	2.8	748	22	AAW46716	D. melanogaster DN
703	6	2.8	517	22	ABP35577	Fungal ZAC protein	776	6	2.8	749	23	ABB04868	LDL receptor bindi
704	6	2.8	518	22	AAU36880	Staphylococcus aur	777	6	2.8	749	23	ABB04869	LDL receptor bindi
705	6	2.8	519	22	ABBS2490	Escherichia coli P	778	6	2.8	749	23	ABB04870	LDL receptor bindi
706	6	2.8	521	21	AAE17561	Arbidopsis thalia	779	6	2.8	749	23	ABB04871	Human polypeptide
707	6	2.8	522	23	AAO17581	M. catarrhalis MCA1	780	6	2.8	752	22	AAW39146	Arbidopsis thalia
708	6	2.8	525	22	ABR64604	Drosophila melanog	781	6	2.8	753	22	AAE01967	Arbidopsis thalia
709	6	2.8	526	22	AAE12634	Human gene 1 encod	782	6	2.8	755	22	ABG21163	Novel human diagno
710	6	2.8	526	22	AAE93132	Human protein sequ	783	6	2.8	756	21	AAE13394	Rat phospholipase
711	6	2.8	530	21	AAV77471	Human deubiquitina	784	6	2.8	776	22	ABG29124	Novel human diagno
712	6	2.8	530	22	AAE64049	Human deubiquitin	785	6	2.8	783	21	AAE51467	Arbidopsis thalia
713	6	2.8	531	22	ABR60201	Drosophila melanog	786	6	2.8	785	21	AAE51466	Arbidopsis thalia
714	6	2.8	537	20	AAW93431	A. thaliana EL5 pr	787	6	2.8	792	22	ABG14278	Novel human diagno
715	6	2.8	541	22	ABR62565	Drosophila melanog	788	6	2.8	793	21	AAE51465	Arbidopsis thalia
716	6	2.8	545	22	AAW40860	Human polypeptide	789	6	2.8	795	23	AAW50347	Variant of Lu-ECAM
717	6	2.8	553	21	AAV97004	S. cerevisiae esse	790	6	2.8	796	23	ABG61593	Human DPPIV relate
718	6	2.8	557	23	ABR90291	Human polypeptide	791	6	2.8	796	23	ABB04588	Human antinopeptida
719	6	2.8	557	23	AAU75086	Ryegrass 4-coumar	792	6	2.8	799	22	ABBS8486	Drosophila melanog
720	6	2.8	559	22	AAU36493	Pseudomonas aerugi	793	6	2.8	803	22	ABBS9443	Drosophila melanog
721	6	2.8	563	23	AAU74557	Human kinesin moto	794	6	2.8	804	18	AAW1387	Fission yeast prot
722	6	2.8	564	23	ABP30223	Streptococcus poly	795	6	2.8	805	22	AAU47293	Murine Ah receptor
723	6	2.8	567	23	ABR25544	Streptococcus poly	796	6	2.8	805	22	AAU47293	Proprionbacterium
724	6	2.8	572	23	ABP27902	Streptococcus poly	797	6	2.8	810	23	ABP35711	Fungal ZBC protein
725	6	2.8	586	20	AAV23622	Protein encoded by	798	6	2.8	812	22	AAW40932	Human polypeptide
726	6	2.8	587	21	AAW44549	Violence gene pro	799	6	2.8	816	16	AAE66931	AMML chromosome in
727	6	2.8	587	21	AAU41035	Propionibacterium	800	6	2.8	821	23	AAW50348	Variant of Lu-ECAM
728	6	2.8	588	22	AAW96513	Putative P. abyssi	801	6	2.8	821	23	ABW47412	Listeria monocytog
729	6	2.8	588	22	AAW39597	Human polypeptide	802	6	2.8	822	20	AAV35538	Amino acid sequenc
730	6	2.8	588	23	AAW49551	Actinoplanes sp ac	803	6	2.8	823	22	ABBS61359	Drosophila melanog
731	6	2.8	590	22	AAE90229	C. glutamicum prote	804	6	2.8	825	22	ABBS6507	Drosophila melanog
732	6	2.8	593	22	AAE93601	Human protein sequ	805	6	2.8	833	21	AAE12554	Human ovarian carc
733	6	2.8	598	22	AAE80043	Corynebacterium gl	806	6	2.8	833	23	ABP30898	0772p clone 21003.
734	6	2.8	612	22	AAE04787	Vigna unguiculata	807	6	2.8	836	21	AAE042123	Arbidopsis thalia
735	6	2.8	613	22	ABG08163	Novel human diagno	808	6	2.8	836	22	AAE01969	Arbidopsis thalia
736	6	2.8	621	22	ABBS8678	Drosophila melanog	809	6	2.8	841	22	AAE01968	Arbidopsis thalia
737	6	2.8	622	22	AAW39596	Human polypeptide	810	6	2.8	843	22	ABBS6933	Drosophila melanog
738	6	2.8	627	20	AAW85596	Human GABA-A recep	811	6	2.8	852	20	AAV30948	Human E3 ubiquitin
739	6	2.8	627	20	AAW85597	Human GABA-A recep	812	6	2.8	854	20	AAV30949	Murine E3 ubiquitin
740	6	2.8	629	22	AAU32412	Novel human secret	813	6	2.8	854	21	AAV95074	Candida albicans p

814	6	2.8	856	21	AA642122	Arabidopsis thalia	887	6	2.8	1468	22	AB565329	Drosophila melanog
815	6	2.8	859	21	AA96997	S. cerevisiae esee	888	6	2.8	1472	22	AB562283	Drosophila melanog
816	6	2.8	884	21	AB891405	Hericidially activ	889	6	2.8	1527	22	AB557771	Drosophila melanog
817	6	2.8	885	16	AA669930	AML chromosome in	890	6	2.8	1544	20	AA441109	Human cancer-associ
818	6	2.8	893	21	AA95012	Novel human diago	891	6	2.8	1554	22	AB111890	Human transcriptio
819	6	2.8	899	22	AB527691	Novel human diago	892	6	2.8	1603	23	AA514685	Human transcriptio
820	6	2.8	904	22	AB562159	Drosophila melanog	893	6	2.8	1648	23	AB554925	Lactococcus lactis
821	6	2.8	904	22	AB562853	Drosophila melanog	894	6	2.8	1651	23	AB566725	Human novel polype
822	6	2.8	905	23	AA505345	Bovine lung endoth	895	6	2.8	1654	22	AB565684	Drosophila melanog
823	6	2.8	909	21	AA642121	Arabidopsis thalia	896	6	2.8	1674	23	AB591548	Hericidially activ
824	6	2.8	911	19	AAW74723	Human secreted pro	897	6	2.8	1675	21	AA542658	Human ORFX ORF2422
825	6	2.8	911	23	AB561838	Human polypeptide	898	6	2.8	1755	21	AA533006	Arabidopsis thalia
826	6	2.8	914	21	AB512552	Human ovarian carc	899	6	2.8	1788	20	AA506077	Human actin-bindin
827	6	2.8	914	22	AB599203	Human ovarian tumor	900	6	2.8	1792	20	AA506078	Bovine actin-bindi
828	6	2.8	914	23	AB508936	0772P protein. Ho	901	6	2.8	1793	21	AA533005	Arabidopsis thalia
829	6	2.8	914	23	AB509668	Hypotheetical prote	902	6	2.8	1826	20	AA534695	C. pneumoniae proc
830	6	2.8	920	22	AB566456	Protein encoded by	903	6	2.8	1839	21	AA533004	Arabidopsis thalia
831	6	2.8	921	21	AA532237	Corn polycarb prot	904	6	2.8	1857	23	AA508430	Protein MYH11 dift
832	6	2.8	921	23	AB509711	Sequence of H4p he	905	6	2.8	1879	22	AA525750	Human actin-bindin
833	6	2.8	922	22	AB505049	Human homology of D	906	6	2.8	1935	22	AB559858	Drosophila melanog
834	6	2.8	922	23	AB505708	Human signal trans	907	6	2.8	1969	19	AAW72419	Rice bacterial lea
835	6	2.8	930	23	AB509708	Sequence of H4p he	908	6	2.8	1972	17	AAW00024	Smooth muscle myos
836	6	2.8	930	23	AB509709	Amino acid sequenc	909	6	2.8	2213	20	AA506079	Human actin-bindin
837	6	2.8	931	23	AA014529	Protein of Me21 (M	910	6	2.8	2230	22	AB560066	Drosophila melanog
838	6	2.8	932	23	AB509706	Sequence of H4p he	911	6	2.8	2383	21	AB515945	Novel human diago
839	6	2.8	933	23	AB509707	Novel central nerv	912	6	2.8	2400	22	AB520278	Novel human diago
840	6	2.8	946	22	AA508710	Novel central nerv	913	6	2.8	2415	22	AB558710	Drosophila melanog
841	6	2.8	967	20	AAW93621	Human CD13/aminope	914	6	2.8	2455	22	AB520279	Novel human diago
842	6	2.8	969	23	AB598216	Chlamydia polypep	915	6	2.8	2633	22	AB506505	Novel human diago
843	6	2.8	971	22	AB595599	Drosophila melanog	916	6	2.8	2639	23	AA5075489	S. aureus antigen1
844	6	2.8	977	21	AB545345	Human pancreatic c	917	6	2.8	2663	22	AA539097	Human polypeptide
845	6	2.8	978	22	AB595935	Drosophila melanog	918	6	2.8	2688	22	AA540883	Human polypeptide
846	6	2.8	983	22	AB570130	Drosophila melanog	919	6	2.8	2732	22	AB552855	Escherichia coli p
847	6	2.8	984	23	AB525627	Fungal ZBC protein	920	6	2.8	2837	22	AB557791	Drosophila melanog
848	6	2.8	996	23	AB591788	Hericidially activ	921	6	2.8	2853	19	AA598828	H. pylori GHP 148
849	6	2.8	997	23	AB580927	Mouse BCO2 ortholo	922	6	2.8	2893	19	AAW71556	Helicobacter polyp
850	6	2.8	999	23	AB580926	Human breast cance	923	6	2.8	3010	20	AA520477	HCV-S1 full-length
851	6	2.8	999	23	AB561893	Prostate cancer acti	924	6	2.8	3011	13	AA522154	NAV1N Hutch c59 is
852	6	2.8	1000	23	AA505353	Human calcium acti	925	6	2.8	3011	23	AA519888	Hepatitis C virus
853	6	2.8	1001	22	AB520445	Novel human diago	926	6	2.8	3095	23	AA520788	Rat C3b/C4b comple
854	6	2.8	1014	23	AA017503	C. elegans transmem	927	6	2.8	3298	22	AA503657	Human extracellular
855	6	2.8	1038	22	AB594063	Human protein sequ	928	6	2.8	3451	23	AB531026	Amino acid sequenc
856	6	2.8	1046	23	AB554658	Lactococcus lactis	929	6	2.8	3562	22	AA584213	Polyketide synthas
857	6	2.8	1056	22	AB566146	Drosophila melanog	930	6	2.8	3567	14	AA544431	eryA region polype
858	6	2.8	1056	22	AB569318	Drosophila melanog	931	6	2.8	4095	21	AA549938	A DNA-dependent pr
859	6	2.8	1077	22	AB561324	Drosophila melanog	932	6	2.8	4096	20	AA584372	Drosophila melanog
860	6	2.8	1105	22	AB558500	Drosophila melanog	933	6	2.8	4097	22	AB557815	Drosophila melanog
861	6	2.8	1119	22	AB578825	Thermus aquaticus	934	6	2.8	4152	22	AA584992	Shrimp white spot
862	6	2.8	1119	23	AA514716	Thermus aquaticus	935	6	2.8	4643	22	AA584992	Drosophila melanog
863	6	2.8	1148	22	AB550283	HOSR-1 ovarian tum	936	6	2.8	4899	22	AB565885	Drosophila melanog
864	6	2.8	1148	22	AB558536	Human protein sequ	937	6	2.8	26926	22	AA505396	Human titin (come
865	6	2.8	1148	23	AB509664	Truncated version	938	6	2.8	26926	14	AA534098	Human titin (come
866	6	2.8	1148	23	AB509664	Truncated version	939	6	2.8	26926	14	AA534098	Human titin (come
867	6	2.8	1151	23	AB598357	Staphylococcus epi	940	6	2.8	26926	14	AA534098	Human titin (come
868	6	2.8	1156	23	AB509665	Protein with sequ	941	6	2.8	26926	14	AA534098	Human titin (come
869	6	2.8	1170	23	AB583450	Human kinase. Hom	942	6	2.8	26926	14	AA534098	Human titin (come
870	6	2.8	1174	22	AB561704	Drosophila melanog	943	6	2.8	26926	14	AA534098	Human titin (come
871	6	2.8	1178	21	AA571311	Human neurite grow	944	6	2.8	26926	14	AA534098	Human titin (come
872	6	2.8	1185	19	AAW98508	H. pylori GHP 123	945	6	2.8	26926	14	AA534098	Human titin (come
873	6	2.8	1185	22	AA535873	Helicobacter pylor	946	6	2.8	26926	14	AA534098	Human titin (come
874	6	2.8	1192	21	AA556967	Human MAGI polypep	947	6	2.8	26926	14	AA534098	Human titin (come
875	6	2.8	1192	22	AA504591	Human Nogo protei	948	6	2.8	26926	14	AA534098	Human titin (come
876	6	2.8	1192	22	AA582349	Human Nogo- $\gamma$ prote	949	6	2.8	26926	14	AA534098	Human titin (come
877	6	2.8	1210	22	AA511767	Human kinase (PKIN	950	6	2.8	26926	14	AA534098	Human titin (come
878	6	2.8	1210	23	AA515798	Human protein kin	951	6	2.8	26926	14	AA534098	Human titin (come
879	6	2.8	1246	22	AA533228	Novel human secret	952	6	2.8	26926	14	AA534098	Human titin (come
880	6	2.8	1250	22	AB512554	Human S3-12 homolo	953	6	2.8	26926	14	AA534098	Human titin (come
881	6	2.8	1257	21	AA585281	Human albinism Cadr32	954	6	2.8	26926	14	AA534098	Human titin (come
882	6	2.8	1325	23	AA021705	Human secreted pro	955	6	2.8	26926	14	AA534098	Human titin (come
883	6	2.8	1337	16	AA585203	hudep-1. Homo gap	956	6	2.8	26926	14	AA534098	Human titin (come
884	6	2.8	1362	20	AAW99481	Murine p/Clp prote	957	6	2.8	26926	14	AA534098	Human titin (come
885	6	2.8	1373	23	AB557364	Mouse ischaemic co	958	6	2.8	26926	14	AA534098	Human titin (come
886	6	2.8	1408	19	AAW62871	Multiple drug resi	959	6	2.8	26926	14	AA534098	Human titin (come



960	5	2.3	9	22	ABP24308	HIV A24 motif pol
961	5	2.3	9	22	AAW22408	HIV peptide SEQ ID
962	5	2.3	9	22	AAW23115	HIV peptide SEQ ID
963	5	2.3	9	22	AAW75708	HIV Class I bindin
964	5	2.3	9	23	ABW96944	Human tumour anti
965	5	2.3	9	23	ABW97130	Human tumour anti
966	5	2.3	9	23	AAW80666	Human tumour a-asoc
967	5	2.3	10	17	AAW00682	Peptide comprising
968	5	2.3	10	21	AAW86625	Telomerase peptide
969	5	2.3	10	21	AAW66270	HIV A2-binding HIV
970	5	2.3	10	21	AAW65347	HIV A28-binding HI
971	5	2.3	10	22	ABP13598	HIV A24 super moti
972	5	2.3	10	22	ABP16133	HIV A24 super moti
973	5	2.3	10	22	ABP19451	HIV B62 super moti
974	5	2.3	10	22	ABP23565	HIV A11 motif pol
975	5	2.3	10	22	ABP23566	HIV A11 motif pol
976	5	2.3	10	22	ABW52187	Human APl-214 tryp
977	5	2.3	10	22	AAW43052	Mycoplasma genital
978	5	2.3	10	23	ABW76266	Human lens thioltr
979	5	2.3	10	23	ABW96945	Human tumour anti
980	5	2.3	10	23	ABW97131	Human tumour anti
981	5	2.3	11	16	AAW80118	ICAM CD54 (38-48)
982	5	2.3	11	20	AAW74271	HT loop peptide KI
983	5	2.3	11	20	ABP14818	HIV A03 super moti
984	5	2.3	11	22	ABP21450	HIV A03 motif pol
985	5	2.3	11	22	ABP21696	HIV A03 motif pol
986	5	2.3	11	22	ABP21697	HIV A03 motif pol
987	5	2.3	11	22	ABP23369	HIV A11 motif pol
988	5	2.3	11	22	ABP23570	HIV A11 motif pol
989	5	2.3	11	22	ABP23571	HIV A11 motif pol
990	5	2.3	11	22	ABP23576	HIV A11 motif pol
991	5	2.3	11	22	ABP23577	HIV A11 motif pol
992	5	2.3	11	23	ABW67561	Human ADP1 tryptic
993	5	2.3	12	16	AAW80117	ICAM CD54 (37-48)
994	5	2.3	12	16	AAW80116	ICAM CD54 (37-48)
995	5	2.3	12	18	AAW34098	ICAM-1 Delta subun
996	5	2.3	12	18	AAW34101	ICAM-1 Delta subun
997	5	2.3	12	21	AAW13059	Cyclic peptide use
998	5	2.3	12	22	ABP21319	HIV A03 motif pol
999	5	2.3	12	22	ABP21430	HIV A03 motif pol
1000	5	2.3	12	22	ABP21453	HIV A03 motif pol

## ALIGNMENTS

RESULT 1  
ABG62118 standard; Protein; 85 AA.

XX  
AC ABG62118;  
XX  
DT 19-AUG-2002 (first entry)  
XX  
DE Human prostate specific polypeptide #1.  
XX  
KW Prostate specific polypeptide; metastasis; prostate cancer; cancer;  
KW non-cancerous prostate disease; gene therapy.  
XX  
OS Homo sapiens.  
XX  
PN WO200238810-A2.  
XX  
PD 16-MAY-2002.  
XX  
PF 06-NOV-2001; 2001WO-US47001.  
XX  
PR 06-NOV-2000; 2000US-246109P.  
XX  
PA (DIAD-) DIADEXUS INC.  
XX  
PI Sun Y, Recipon H, Chen S, Liu C;  
XX

DR WPI; 2002-471559/50.  
XX  
PT New prostate specific polypeptides and polynucleotides useful for  
PT detecting, diagnosing, monitoring, treating, staging and predicting  
PT cancers in humans having cancer and non-cancerous prostate disease  
XX  
XX  
XX Claim 11; Page 242; 267pp; English.  
XX  
XX The invention describes an isolated prostate specific polypeptide (I) and  
XX nucleic acid (II) encoding it and are useful for diagnosing and  
XX monitoring the presence and metastases of prostate cancer in a patient.  
XX (I), (II) and an antibody to (II) are useful in quantitative and  
XX qualitative diagnostic assays and methods for detecting, diagnosing,  
XX monitoring, treating, staging and predicting cancers in humans having  
XX cancer or may have the risk of developing cancer. (I) and (II) are also  
XX useful for determining non-cancerous prostate disease, by measuring  
XX their expression levels and/or structural alterations; for determining  
XX the sample that has prostate tissue-like characteristics or is a prostate  
XX tissue; as an element in an array or a multi-analyte test to recognise  
XX expression patterns associated with prostate cancer and other prostate  
XX related disorders; and as elements in a computer program for pattern  
XX recognition of prostate disorders. (I) and (II) are useful for producing  
XX engineered prostate tissue for treatment and research. (II) is useful for  
XX producing transgenic animals and cells and also in gene therapy. This  
XX is the amino acid sequence of the prostate specific nucleic acid  
XX described in the invention.

Sequence 85 AA:

Query Match 3.7%; Score 8; DB 23; Length 85;  
Best local Similarity 100.0%; Pred. No. 3.3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 VREPKIEL 38  
Db 27 VREPKIEL 34

RESULT 2  
ABP33399 standard; Protein; 74 AA.

XX  
AC ABP33399;  
XX  
DT 08-JUL-2002 (first entry)  
XX  
DE Human ORF2372 protein, SEQ ID NO:4744.  
XX  
XX  
XX Human; ORF; open reading frame; ORF; drug screening; diagnosis;  
XX disease monitoring; cytokine; cell proliferation; cell differentiation;  
XX immune modulation; haematopoiesis regulation; tissue growth;  
XX angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;  
XX thrombolytic; tumour inhibition; bodily characteristics; fertility;  
XX behaviour; cancer; proliferative disorder; neurological disorder;  
XX cardiovascular disease; immune system disorder; organ transplantation;  
XX tissue growth disorder; tissue regeneration disorder; diabetes mellitus;  
XX hypothyroidism; cholesterol ester storage disease; infection; vulvovag;  
XX vasotropic; antipsoriatic; antidiabetic; cytostatic; nootropic;  
XX neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;  
XX cardiant; hypotensive; antihypertoid; antiinflammatory; immunomodulator;  
XX dermatological; analgesic; virucide; antibacterial; fungicide.  
XX  
XX Homo sapiens.  
XX  
XX OS  
XX PN WO200190366-A2.  
XX  
XX PD 29-NOV-2001.  
XX  
XX PF 24-MAY-2001; 2001WO-US17076.  
XX  
XX PR 24-MAY-2000; 2000US-206690P.  
XX  
XX PA (CURA-) CURAGEN CORP.  
XX

XX Leach MD, Shinkete RA;  
PI WPI; 2002-106200/14.  
DR N-PSDB; ABN77425.  
XX  
PT Novel human polypeptides and polynucleotides useful for diagnosing,  
PT preventing and treating cardiovascular disease, neurodegenerative,  
PT hyperproliferative disorders and disorders related to organ  
PT transplantation -  
XX  
XX Claim 10, Page 1439; 2508pp; English.  
XX  
XX Sequences ABP31028-ABP3561 represent 4534 novel human proteins  
CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-  
CC ABN79587 represent cDNAs encoding them. The invention also encompasses  
CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively  
CC referred to as ORFX) proteins, polynucleotides at least 85% identical to  
CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX  
CC polynucleotides, the recombinant production of ORFX proteins, antibodies  
CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and  
CC polypeptides, methods of screening for modulators of ORFX expression or  
CC activity, and methods of screening individuals for a predisposition to an  
CC ORFX-associated disorder. The ORFX proteins of the invention have a wide  
CC range of biological activities, such as cytokine, cell proliferation,  
CC cell differentiation, immune modulation, haematopoiesis regulation,  
CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/  
CC chemokinetic activity, haemostatic activity, thrombolytic activity,  
CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,  
CC and antiinfective activity, and may also be involved in the determination  
CC of bodily characteristics, fertility and behaviour. ORFX proteins,  
CC nucleic acids and antibodies may be used in the treatment of cancers,  
CC other proliferative disorders such as psoriasis and benign tumours,  
CC neurological disorders such as epilepsy and Alzheimer's disease,  
CC cardiovascular diseases, immune system disorders, disorders related to  
CC organ transplantation, disorders of tissue growth and regeneration,  
CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester  
CC storage disease, and infectious diseases caused by viral, bacterial,  
CC fungal and other pathogens. ORFX nucleic acids may also be used as a  
CC source of primers and probes, in the detection of ORFX genomic sequences  
CC or transcripts, in the identification and cloning of homologous  
CC sequences, in genetic diagnosis, and in forensic biology. The ORFX  
CC nucleic acids may additionally be used to produce transgenic animals  
CC which may be useful for studying the function and/or activity of ORFX  
CC protein, and in drug screening. The ORFX proteins may also be used as  
CC immunogens to generate specific antibodies, which are useful in the  
CC diagnosis, treatment and monitoring of ORFX-associated diseases.  
XX  
XX Sequence 74 AA;  
SQ  
Query Match 3.2%; Score 7; DB 23; Length 74;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 170 LNSGSVL 176  
|||  
Db 66 LNSGSVL 72  
RESULT 3  
AAG12854  
ID AAG12854 standard; Protein; 139 AA.  
XX  
XX AAG12854;  
AC  
XX  
XX 17-OCT-2000 (first entry)  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 12126.  
XX  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
KM termination sequence.  
XX

OS Arabidopsis thaliana.  
XX  
XX EP1033405-A2.  
XX  
XX 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128214.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134321.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.

```

PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0141844.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142809.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143524.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151086.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.

```

```

PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159295.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 3.2%; Score 7; DB 21; Length 139;  
 Beel Local Similarity 100.0%; Pred. No. 54;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 ADDAEGK 161  
 Db 74 ADDAEGK 80

RESULT 4  
 AAG12853  
 ID AAG12853 standard; Protein. 141 AA.

AAAG12853;  
 17-OCT-2000 (first entry)

Arabidopsis thaliana protein fragment SEQ ID NO: 12125.

Protein identification; signal transduction pathway; metabolic pathway;  
 hybridisation assay; genetic mapping; gene expression control; promoter;  
 termination sequence.

Arabidopsis thaliana.

EP1033405-A2.

06-SEP-2000.

25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126254.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128274.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139482.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140354.  
PR 26-JUN-1999; 99US-0140695.  
PR 29-JUN-1999; 99US-0140823.  
PR 30-JUN-1999; 99US-0140991.  
PR 01-JUL-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142399.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144335.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144332.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145593.  
PR 27-JUL-1999; 99US-0145591.  
PR 27-JUL-1999; 99US-0145599.  
PR 28-JUL-1999; 99US-0145911.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147432.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149375.  
PR 18-AUG-1999; 99US-0149426.  
PR 18-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 25-AUG-1999; 99US-0150566.  
PR 25-AUG-1999; 99US-0150884.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.

PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161921.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 3.2%; Score 7; DB 21; Length 141;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 155 ADDAEGK 161  
|||  
DB 76 ADDAEGK 82

RESULT 5  
AAW20201  
ID AAW20201 standard; Protein; 146 AA.

AC AAW20201;  
XX  
DT 09-JUL-1997 (first entry)  
XX  
DE H. pylori secreted or periplasmic protein, 20415937.aa.  
XX  
KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;  
KW identification; binding compound; bacterium; life cycle; activator;  
KW bacteria; inhibitor; duodenal ulcer disease; chronic gastritis;  
KW diagnosis.

OS Helicobacter pylori.  
XX  
PN WO9640893-A1.  
XX  
PD 19-DEC-1996.  
XX  
PF 06-JUN-1996; 96WO-US09122.  
XX  
PR 01-APR-1996; 96US-0630405.  
XX  
PR 07-JUN-1995; 95US-0487032.  
XX  
PA (ASTRA ) ASTRA AB.  
XX

PI Berglindh OT, Smith D, Welligaerd BL,  
XX  
DR WPI; 1997-052306/05.  
DR N-ESDB; AAT67431.  
XX  
PT Helicobacter pylori nucleic acid sequences and related  
PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori  
PT infection, and to detect Helicobacter  
XX  
PS Claim 72; Page 415-16; 1481pp; English.  
XX  
CC This sequence represents a H. pylori secreted or periplasmic protein.  
CC The protein may be used in a vaccine to prevent or treat H. pylori  
CC infection or to identify H. pylori polypeptide binding compounds,  
CC useful as potential H. pylori life cycle activators or inhibitors.  
CC The genomic sequence of H. pylori (ATCC 55679) was determined from  
CC overlapping contigs generated by mechanically shearing the bacterial  
CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,  
CC and the predicted coding regions defined by computer evaluation. To  
CC identify likely H. pylori antigens for vaccine development, the amino  
CC acid sequences predicted from various ORF were analysed for significant  
CC homology to other known or exported membrane proteins. Having identified  
CC and determined the sequences of interest, particular regions can be  
CC isolated from H. pylori by PCR amplification for recombinant polypeptide  
CC production, e.g. in E. coli hosts.  
XX  
SQ Sequence 146 AA;

Query Match 3.2%; Score 7; DB 18; Length 146;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 EPKIELA 39  
|||  
DB 105 EPKIELA 111

RESULT 6  
AAG12852  
ID AAG12852 standard; Protein; 148 AA.

AC AAG12852;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 12124.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.

OS Arabidopsis thaliana.

XX  
PN EP1033405-A2.

XX  
PD 06-SEP-2000.

XX  
PF 25-FEB-2000; 2000EP-0301439.

XX  
PR 25-FEB-1999; 99US-0121825.  
XX  
PR 05-MAR-1999; 99US-0121180.  
XX  
PR 09-MAR-1999; 99US-0123548.  
XX  
PR 23-MAR-1999; 99US-0125788.  
XX  
PR 25-MAR-1999; 99US-0126264.  
XX  
PR 29-MAR-1999; 99US-0126785.  
XX  
PR 01-APR-1999; 99US-0127462.  
XX  
PR 06-APR-1999; 99US-0128234.  
XX  
PR 08-APR-1999; 99US-0128714.  
XX  
PR 16-APR-1999; 99US-0129845.  
XX  
PR 19-APR-1999; 99US-0130077.  
XX  
PR 21-APR-1999; 99US-0130449.  
XX  
PR 23-APR-1999; 99US-0130510.  
XX  
PR 23-APR-1999; 99US-0130891.

---

PR 28-APR-1999;	99US-011449.	PR 21-JUL-1999;	99US-0144814.
PR 30-APR-1999;	99US-0132048.	PR 21-JUL-1999;	99US-0145086.
PR 04-MAY-1999;	99US-0132407.	PR 21-JUL-1999;	99US-0145088.
PR 05-MAY-1999;	99US-0132484.	PR 22-JUL-1999;	99US-0145085.
PR 06-MAY-1999;	99US-0132485.	PR 22-JUL-1999;	99US-0145087.
PR 06-MAY-1999;	99US-0132486.	PR 22-JUL-1999;	99US-0145089.
PR 07-MAY-1999;	99US-0132487.	PR 22-JUL-1999;	99US-0145192.
PR 11-MAY-1999;	99US-0132863.	PR 23-JUL-1999;	99US-0145145.
PR 14-MAY-1999;	99US-0134218.	PR 23-JUL-1999;	99US-0145218.
PR 14-MAY-1999;	99US-0134219.	PR 26-JUL-1999;	99US-0145276.
PR 14-MAY-1999;	99US-0134221.	PR 27-JUL-1999;	99US-0145913.
PR 14-MAY-1999;	99US-0134370.	PR 27-JUL-1999;	99US-0145918.
PR 18-MAY-1999;	99US-0134768.	PR 27-JUL-1999;	99US-0145919.
PR 19-MAY-1999;	99US-0134941.	PR 28-JUL-1999;	99US-0145951.
PR 20-MAY-1999;	99US-0135124.	PR 02-AUG-1999;	99US-0146386.
PR 21-MAY-1999;	99US-0135353.	PR 02-AUG-1999;	99US-0146388.
PR 24-MAY-1999;	99US-0135629.	PR 02-AUG-1999;	99US-0146389.
PR 25-MAY-1999;	99US-0136021.	PR 03-AUG-1999;	99US-0147038.
PR 27-MAY-1999;	99US-0136392.	PR 04-AUG-1999;	99US-0147204.
PR 28-MAY-1999;	99US-0136782.	PR 04-AUG-1999;	99US-0147302.
PR 01-JUN-1999;	99US-0137222.	PR 05-AUG-1999;	99US-0147392.
PR 03-JUN-1999;	99US-0137528.	PR 05-AUG-1999;	99US-0147260.
PR 04-JUN-1999;	99US-0137502.	PR 06-AUG-1999;	99US-0147303.
PR 07-JUN-1999;	99US-0137724.	PR 06-AUG-1999;	99US-0147416.
PR 08-JUN-1999;	99US-0138094.	PR 09-AUG-1999;	99US-0147493.
PR 10-JUN-1999;	99US-0138540.	PR 10-AUG-1999;	99US-0147935.
PR 10-JUN-1999;	99US-0138847.	PR 11-AUG-1999;	99US-0148171.
PR 14-JUN-1999;	99US-0139119.	PR 12-AUG-1999;	99US-0148319.
PR 16-JUN-1999;	99US-0139452.	PR 13-AUG-1999;	99US-0148341.
PR 16-JUN-1999;	99US-0139453.	PR 13-AUG-1999;	99US-0148565.
PR 17-JUN-1999;	99US-0139492.	PR 16-AUG-1999;	99US-0148684.
PR 18-JUN-1999;	99US-0139454.	PR 17-AUG-1999;	99US-0149368.
PR 18-JUN-1999;	99US-0139455.	PR 18-AUG-1999;	99US-0149175.
PR 18-JUN-1999;	99US-0139456.	PR 20-AUG-1999;	99US-0149426.
PR 18-JUN-1999;	99US-0139457.	PR 20-AUG-1999;	99US-0149722.
PR 18-JUN-1999;	99US-0139458.	PR 20-AUG-1999;	99US-0149723.
PR 18-JUN-1999;	99US-0139459.	PR 23-AUG-1999;	99US-0149929.
PR 18-JUN-1999;	99US-0139460.	PR 23-AUG-1999;	99US-0149902.
PR 18-JUN-1999;	99US-0139461.	PR 25-AUG-1999;	99US-0149930.
PR 18-JUN-1999;	99US-0139462.	PR 26-AUG-1999;	99US-0150566.
PR 18-JUN-1999;	99US-0139463.	PR 26-AUG-1999;	99US-0150884.
PR 18-JUN-1999;	99US-0139750.	PR 27-AUG-1999;	99US-0151065.
PR 18-JUN-1999;	99US-0139763.	PR 27-AUG-1999;	99US-0151066.
PR 21-JUN-1999;	99US-0139817.	PR 27-AUG-1999;	99US-0151080.
PR 22-JUN-1999;	99US-0139891.	PR 30-AUG-1999;	99US-0151303.
PR 23-JUN-1999;	99US-0140353.	PR 31-AUG-1999;	99US-0151438.
PR 23-JUN-1999;	99US-0140354.	PR 01-SEP-1999;	99US-0151930.
PR 24-JUN-1999;	99US-0140695.	PR 07-SEP-1999;	99US-0152363.
PR 28-JUN-1999;	99US-0140823.	PR 10-SEP-1999;	99US-0153070.
PR 29-JUN-1999;	99US-0140991.	PR 13-SEP-1999;	99US-0153758.
PR 30-JUN-1999;	99US-0141287.	PR 15-SEP-1999;	99US-0154018.
PR 01-JUL-1999;	99US-0141842.	PR 16-SEP-1999;	99US-0154039.
PR 01-JUL-1999;	99US-0142154.	PR 20-SEP-1999;	99US-0154779.
PR 02-JUL-1999;	99US-0142055.	PR 22-SEP-1999;	99US-0155139.
PR 06-JUL-1999;	99US-0142390.	PR 23-SEP-1999;	99US-0155486.
PR 08-JUL-1999;	99US-0142803.	PR 24-SEP-1999;	99US-0156559.
PR 09-JUL-1999;	99US-0142920.	PR 28-SEP-1999;	99US-0156458.
PR 12-JUL-1999;	99US-0142977.	PR 29-SEP-1999;	99US-0156596.
PR 13-JUL-1999;	99US-0143542.	PR 04-OCT-1999;	99US-0157117.
PR 14-JUL-1999;	99US-0143624.	PR 05-OCT-1999;	99US-0157753.
PR 15-JUL-1999;	99US-0144005.	PR 06-OCT-1999;	99US-0157865.
PR 16-JUL-1999;	99US-0144086.	PR 07-OCT-1999;	99US-0158029.
PR 16-JUL-1999;	99US-0144325.	PR 08-OCT-1999;	99US-0158232.
PR 19-JUL-1999;	99US-0144331.	PR 12-OCT-1999;	99US-0158369.
PR 19-JUL-1999;	99US-0144332.	PR 13-OCT-1999;	99US-0159293.
PR 19-JUL-1999;	99US-0144333.	PR 13-OCT-1999;	99US-0159295.
PR 19-JUL-1999;	99US-0144334.	PR 14-OCT-1999;	99US-0159330.
PR 19-JUL-1999;	99US-0144335.	PR 14-OCT-1999;	99US-0159331.
PR 20-JUL-1999;	99US-0144352.	PR 14-OCT-1999;	99US-0159637.
PR 20-JUL-1999;	99US-0144632.	PR 14-OCT-1999;	99US-0159638.
PR 20-JUL-1999;	99US-0144884.		

PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161921.  
PR 28-OCT-1999; 99US-0161922.  
PR 28-OCT-1999; 99US-0161923.  
PR 29-OCT-1999; 99US-0162142.

Query Match 3.2%; Score 7; DB 21; Length 148;  
Best Local Similarity 100.0%; Pred. No. 57;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 155 ADDAEGK 161  
DB 83 ADDAEGK 89

RESULT 7  
AAB25232  
ID AAB25232 standard; Protein; 151 AA.

AC AAB25232;  
DT 27-NOV-2000 (first entry)

DE Eucalyptus grandis cell signalling involved protein seq ID NO:551.

KM Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;  
KM plant cell signalling; modulation; transgenic plant; pathogen; growth;  
KM environmental change; development; cell proliferation; differentiation;  
KM elongation; survival; disease resistance; nutrient metabolism.

OS Eucalyptus grandis.

XX WO200042171-A1.

PD 20-JUL-2000.

PF 11-JAN-2000; 2000MO-US00724.

XX 12-JAN-1999; 99US-0228986.

PR 01-NOV-1999; 99US-0162866.

XX (GENE-) GENESIS RES & DEV CORP LTD.

XX Strabala TJ, Nieuwenhuizen NJ;

XX WPI; 2000-476052/41.

PT Isolated polynucleotide encoding a polypeptide involved in cell

PT signaling used for generating transgenic plants with modified responses

XX to external signals -

PS Claim 3; Page 246; 527pp; English.

CC AAA79263 to AAA79736 and AAB25100 to AAB25570 represent polynucleotide  
CC and protein sequences isolated from eucalyptus (Eucalyptus grandis) or  
CC pine (Pinus radiata also known as Monterey pine). The protein sequences  
CC are involved in cell signalling. The polynucleotide and protein  
CC sequences can be used to modify the response of plant cells to external

CC signals e.g. environmental changes or pathogens during the growth and  
CC development of a plant. They can be used to modify cell proliferation,  
CC differentiation, elongation and survival, resistance to disease and  
CC nutrient metabolism. Examples of modifications which can be produced are  
CC altered fruit ripening and senescence of leaves and flowers e.g. to  
CC delay senescence and prolong the life of cut flowers or enhance  
CC senescence of reproductive organs to engineer sterile plants. Other  
CC modifications can be used to delay senescence in selected cell types or  
CC organs providing fruit and vegetables which have a longer shelf life  
CC between harvest and consumption, or to decrease branching frequency in  
CC forest tree species giving long stretches of valuable knot-free clear  
CC wood which can be used in solid timber furniture and veneers.

SO Sequence 151 AA;

Query Match 3.2%; Score 7; DB 21; Length 151;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 39 AKEAILK 45  
DB 52 AKEAILK 58

RESULT 8  
AAM49083  
ID AAM49083 standard; Protein; 157 AA.

AC AAM49083;

DT 07-MAY-2002 (first entry)

DE Human transcriptional regulation mediator protein hMedp.

KM Human; transcriptional regulation mediator; hMED6; hMedp;

KM RNA polymerase II transcriptional regulation mediator homologue;

KM Yeast Med6 homologue.

XX Homo sapiens.

XX KR99015342-A.

XX 05-MAR-1999.

PF 05-AUG-1997; 97KR-0037392.

XX 05-AUG-1997; 97KR-0037392.

PA (SMSU ) SAMSUNG ELECTRONICS CO LTD.

PI Kim YJ, Lee YC, Min SY, Kim BS;

XX WPI; 2000-220024/19.

DR N-PSDB; ABA96840.

PS Claim 6; Page 16; 29pp; Korean.

CC The invention relates to a human transcriptional regulation mediator

CC gene designated hMED6 and its encoded protein, hMedp. hMedp is a

CC homologue of the Yeast RNA polymerase II transcriptional regulation

CC mediator Med6 (AAM49084, also given in GenBank accession number U78080).

CC The invention also relates to recombinant microorganisms comprising the

CC hMED6 gene. The present sequence represents hMedp.

XX Sequence 157 AA;

Query Match 3.2%; Score 7; DB 21; Length 157;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 170 LNSGSVL 176  
|||||  
Db 24 LNSGSVL 30

## RESULT 9

ABG17907  
ID ABG17907 standard; Protein, 159 AA.

XX  
AC ABG17907;

XX  
DT 18-FEB-2002 (first entry)

XX  
DE Novel human diagnostic protein #17898.

XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX  
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX  
OS Homo sapiens.

XX  
PN MO200175067-A2.

XX  
PD 11-OCT-2001.

XX  
PF 30-MAR-2001; 2001WO-US08631.

XX  
PR 31-MAR-2000; 2000US-0540217.

XX  
PR 23-AUG-2000; 2000US-0649167.

XX  
PA (HXSE-) HXSEQ INC.

XX  
PI Drmanac RT, Liu C, Tang YT;

XX  
DR WPI: 2001-639362/73.

XX  
DR N-PSDB: AAS82094.

PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -

PS Claim 20; SEQ ID No 48266; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and  
XX polypeptide (II) sequences. (I) is useful as hybridization probes,  
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
XX and gene mapping, and in recombinant production of (II). The  
XX polynucleotides are also used in diagnostics as expressed sequence tags  
XX for identifying expressed genes. (I) is useful in gene therapy techniques  
XX to restore normal activity of (II) or to treat disease states involving  
XX (II). (II) is useful for generating antibodies against it, detecting or  
XX quantitating a polypeptide in tissue, as molecular weight markers and as  
XX a food supplement. (II) and its binding partners are useful in medical  
XX imaging of sites expressing (II). (I) and (II) are useful for treating  
XX disorders involving aberrant protein expression or biological activity.  
XX The polypeptide and polynucleotide sequences have applications in  
XX diagnostics, forensics, gene mapping, identification of mutations in  
XX responsible for genetic disorders or other traits to assess biodiversity  
XX and to produce other types of data and products dependent on DNA and  
XX amino acid sequences. ABG0010-ABG30377 represent novel human  
XX diagnostic amino acid sequences of the invention.  
XX Note: The sequence data for this patent did not appear in the printed  
XX specification, but was obtained in electronic format directly from WIRO  
XX at ftp.wipo.int/pub/published\_pct\_sequences.

XX  
XX Sequence 159 AA;

Query Match 3.2%; Score 7; DB 22; Length 159;

Best Local Similarity 100.0%; Pred. No. 61;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 117 AAVILLT 123  
|||||

Db 53 AAVILLT 59

## RESULT 10

AAG12701  
ID AAG12701 standard; Protein, 168 AA.

XX  
AC AAG12701;

XX  
DT 17-OCT-2000 (first entry)

XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 11917.

XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.

XX  
OS Arabidopsis thaliana.

XX  
PN EP1033405-A2.

XX  
PD 06-SEP-2000.

XX  
PF 25-FEB-2000; 2000EP-0301439.

XX  
PR 25-FEB-1999; 99US-0121825.

XX  
PR 05-MAR-1999; 99US-0123180.

XX  
PR 09-MAR-1999; 99US-0123548.

XX  
PR 23-MAR-1999; 99US-0125784.

XX  
PR 25-MAR-1999; 99US-0126264.

XX  
PR 29-MAR-1999; 99US-0126785.

XX  
PR 01-APR-1999; 99US-0127462.

XX  
PR 06-APR-1999; 99US-0128234.

XX  
PR 08-APR-1999; 99US-0128714.

XX  
PR 16-APR-1999; 99US-0129845.

XX  
PR 19-APR-1999; 99US-0130077.

XX  
PR 21-APR-1999; 99US-0130449.

XX  
PR 23-APR-1999; 99US-0130510.

XX  
PR 23-APR-1999; 99US-0130891.

XX  
PR 28-APR-1999; 99US-0131449.

XX  
PR 30-APR-1999; 99US-0132048.

XX  
PR 30-APR-1999; 99US-0132407.

XX  
PR 04-MAY-1999; 99US-0132484.

XX  
PR 05-MAY-1999; 99US-0132485.

XX  
PR 06-MAY-1999; 99US-0132486.

XX  
PR 07-MAY-1999; 99US-0132487.

XX  
PR 11-MAY-1999; 99US-0134256.

XX  
PR 14-MAY-1999; 99US-0134218.

XX  
PR 14-MAY-1999; 99US-0134219.

XX  
PR 14-MAY-1999; 99US-0134370.

XX  
PR 18-MAY-1999; 99US-0134768.

XX  
PR 19-MAY-1999; 99US-0134941.

XX  
PR 20-MAY-1999; 99US-0135124.

XX  
PR 21-MAY-1999; 99US-0135353.

XX  
PR 24-MAY-1999; 99US-0135629.

XX  
PR 25-MAY-1999; 99US-0136021.

XX  
PR 27-MAY-1999; 99US-0136392.

XX  
PR 28-MAY-1999; 99US-0136782.

XX  
PR 01-JUN-1999; 99US-0137222.

XX  
PR 03-JUN-1999; 99US-0137528.

XX  
PR 04-JUN-1999; 99US-0137528.

XX  
PR 07-JUN-1999; 99US-0137724.

XX  
PR 08-JUN-1999; 99US-0138094.

XX  
PR 10-JUN-1999; 99US-0138540.

XX  
PR 14-JUN-1999; 99US-0138847.

XX  
PR 15-JUN-1999; 99US-0139119.

XX  
PR 16-JUN-1999; 99US-0139452.

XX  
PR 17-JUN-1999; 99US-0139453.

XX  
PR 18-JUN-1999; 99US-0139454.

XX  
PR 18-JUN-1999; 99US-0139454.

XX  
PR 18-JUN-1999; 99US-0139455.



PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139761.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146387.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.

PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160776.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161922.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 3.2% Score 7; DB 21; Length 166;  
Best Local Similarity 100.0%; Pred. No. 64;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 ADDAEGK 161  
|||||

DB 74 ADDAEGK 80

RESULT 11  
AAG12700  
ID AAG12700 standard; Protein; 170 AA.  
XX

AC AAG12700;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 11916.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EPI033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126284.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134286.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 18-MAY-1999; 99US-0134370.  
PR 19-MAY-1999; 99US-0134768.  
PR 20-MAY-1999; 99US-0134941.  
PR 21-MAY-1999; 99US-0135174.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137508.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144362.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.

```

PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158025.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match 3.2%; Score 7; DB 21; Length 170;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 155 ADDAEGK 161
   |||||
DB 76 ADDAEGK 82

```

```

RESULT 12
AAR62784
ID AAR62784 standard; Protein; 175 AA.
XX
AC AAR62784;
XX
DT 25-MAY-1995 (first entry)
XX
DE Borrelia M57 antigen vaccine.
XX
KW OspC antigen; vaccine; Lyme disease; borreliosis; immunogen;

```

```

KW serovar typing; restriction fragment length polymorphism;
KW RFLP; Pichia pastoris.
XX
OS Borrelia burgdorferi M57.
XX
PN WO9425596-A.
XX
PD 10-NOV-1994.
XX
PE 29-APR-1994; 94WO-EP01365.
XX
PR 29-APR-1993; 93US-0053863.
XX
PA (IMMO ) IMMUNO AG.
XX
PI Crowe B, Dornier F, Lavey I;
XX
DR WPI; 1994-358273/44.
XX
DR N-PSDB; AAQ73896.
XX
PT Immunogenic composition comprising OspC antigens - for the
PT treatment of Lyme borreliosis in different, specific geographical
PT areas.
XX
PS Disclosure; Fig. 9; 115pp; English.
XX
CC A vaccine for Lyme disease includes selected OspC antigen
CC formulations based on defined OspC families resolved by serovar
CC typing and RFLP typing. Partial sequences of OspC genes selected
CC from different RFLP types are given in AAQ73883-905 (encoded peptides,
CC completing the first 92% of mature OspC, are given in AAR62771-93).
CC Complete sequences of these novel ospC genes, including the 3' end,
CC plus sequences for the ospC genes of Borrelia strains H13 and 28691
CC are given in AAQ73857-82, and encoded proteins in AAR60884-909. The
CC DNA sequences may be expressed in e.g. Pichia pastoris for
CC recombinant antigen production.
XX
SQ Sequence 175 AA;

```

```

Query Match 3.2%; Score 7; DB 15; Length 175;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 39 AKEAILK 45
   |||||
DB 136 AKEAILK 142

```

```

RESULT 13
AAR62772
ID AAR62772 standard; Protein; 176 AA.
XX
AC AAR62772;
XX
DT 25-MAY-1995 (first entry)
XX
DE Borrelia B31 antigen vaccine.
XX
KW OspC antigen; vaccine; Lyme disease; borreliosis; immunogen;
KW serovar typing; restriction fragment length polymorphism;
KW RFLP; Pichia pastoris.
XX
OS Borrelia burgdorferi B31.
XX
XX
PN WO9425596-A.
XX
PD 10-NOV-1994.
XX
PE 29-APR-1994; 94WO-EP01365.
XX
PR 29-APR-1993; 93US-0053863.
XX
PA (IMMO ) IMMUNO AG.

```

XX Crowe B, Dornier F, Livey I;  
PI WPI; 1994-358273/44.  
XX DR N-PSDB; AA073884.  
XX PT Immunogenic composition comprising OspC antigens - for the  
PT treatment of Lyme borreliosis in different, specific geographical  
PT areas.  
XX PS Disclosure; Fig. 9; 115pp; English.  
XX CC A vaccine for Lyme disease includes selected OspC antigen  
XX CC formulations based on defined OspC families resolved by serovar  
CC typing and RFLP typing. Partial sequences of OspC genes selected  
CC from different RFLP types are given in AA073883-905 (encoded peptides,  
CC comprising the first 92% of mature OspC, are given in AAR62771-93).  
CC Complete sequences of these novel ospc genes, including the 3' end,  
CC plus sequences for the ospc genes of Borrelia strains H13 and 28691  
CC are given in AA073857-82, and encoded proteins in AAR60884-909. The  
CC DNA sequences may be expressed in e.g. Pichia pastoris for  
CC recombinant antigen production.  
XX SQ Sequence 176 AA;  
XX  
XX Query Match 3.2%; Score 7; DB 15; Length 176;  
Best Local Similarity 100.0%; Pred. No. 67;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 39 AKEAILK 45  
Db 137 AKEAILK 143  
Db 137 AKEAILK 143  
RESULT 14  
AAR62785  
ID AAR62785 standard; Protein; 176 AA.  
XX AAR62785;  
XX AC  
XX DT 25-MAY-1995 (first entry)  
XX DE Borrelia W antigen vaccine.  
XX KM OspC antigen; vaccine; Lyme disease; borreliosis; immunogen;  
KM serovar typing; restriction fragment length polymorphism;  
KM RFLP; Pichia pastoris.  
XX OS Borrelia burgdorferi W.  
XX PN WO9425596-A.  
XX PD 10-NOV-1994.  
XX PE 29-APR-1994; 94MO-EP01365.  
XX PR 29-APR-1993; 93US-0053863.  
XX PA (IMMO) IMMUNO AG.  
XX PI Crowe B, Dornier F, Livey I;  
XX WPI; 1994-358273/44.  
XX DR N-PSDB; AA073897.  
XX PT Immunogenic composition comprising OspC antigens - for the  
PT treatment of Lyme borreliosis in different, specific geographical  
PT areas.  
XX PS Disclosure; Fig. 9; 115pp; English.  
XX A vaccine for Lyme disease includes selected OspC antigen  
CC formulations based on defined OspC families resolved by serovar

CC typing and RFLP typing. Partial sequences of OspC genes selected  
CC from different RFLP types are given in AA073883-905 (encoded peptides,  
CC comprising the first 92% of mature OspC, are given in AAR62771-93).  
CC Complete sequences of these novel ospc genes, including the 3' end,  
CC plus sequences for the ospc genes of Borrelia strains H13 and 28691  
CC are given in AA073857-82, and encoded proteins in AAR60884-909. The  
CC DNA sequences may be expressed in e.g. Pichia pastoris for  
CC recombinant antigen production.  
XX SQ Sequence 176 AA;  
XX  
XX Query Match 3.2%; Score 7; DB 15; Length 176;  
Best Local Similarity 100.0%; Pred. No. 67;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 39 AKEAILK 45  
Db 137 AKEAILK 143  
Db 137 AKEAILK 143  
RESULT 15  
AAG12699  
ID AAG12699 standard; Protein; 177 AA.  
XX AAG12699;  
XX AC  
XX DT 17-OCT-2000 (first entry)  
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 11915.  
XX KM Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
KM termination sequence.  
XX OS Arabidopsis thaliana.  
XX PN EP1033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-0301439.  
XX PE 25-FEB-1999; 99US-0121825.  
XX PR 05-MAR-1999; 99US-0123180.  
XX PR 09-MAR-1999; 99US-0123548.  
XX PR 23-MAR-1999; 99US-0125788.  
XX PR 25-MAR-1999; 99US-0126264.  
XX PR 29-MAR-1999; 99US-0127462.  
XX PR 01-APR-1999; 99US-0128234.  
XX PR 06-APR-1999; 99US-0128714.  
XX PR 08-APR-1999; 99US-0129845.  
XX PR 16-APR-1999; 99US-0130077.  
XX PR 19-APR-1999; 99US-0130449.  
XX PR 21-APR-1999; 99US-0130510.  
XX PR 23-APR-1999; 99US-0130891.  
XX PR 28-APR-1999; 99US-0131449.  
XX PR 30-APR-1999; 99US-0132048.  
XX PR 30-APR-1999; 99US-0132407.  
XX PR 04-MAY-1999; 99US-0132484.  
XX PR 05-MAY-1999; 99US-0132485.  
XX PR 06-MAY-1999; 99US-0132486.  
XX PR 07-MAY-1999; 99US-0132487.  
XX PR 11-MAY-1999; 99US-0132863.  
XX PR 14-MAY-1999; 99US-0134256.  
XX PR 14-MAY-1999; 99US-0134219.  
XX PR 14-MAY-1999; 99US-0134221.  
XX PR 14-MAY-1999; 99US-0134370.  
XX PR 18-MAY-1999; 99US-0134768.  
XX PR 19-MAY-1999; 99US-0134941.  
XX PR 20-MAY-1999; 99US-0135124.  
XX PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139452.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142820.  
PR 12-JUL-1999; 99US-0142877.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147312.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149310.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 23-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158365.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.



XX MO9425596-A.  
 PN  
 XX 10-NOV-1994.  
 PD  
 XX  
 XX 29-APR-1994; 94MO-EP01365.  
 PF  
 XX 29-APR-1993; 93US-0053863.  
 PR  
 XX (IMMO ) IMMUNO AG.  
 PA  
 XX Crowe B, Dornier F, Livey I;  
 PI  
 XX WPI; 1994-358273/44.  
 DR  
 DR N-PSDB; AAQ73859.  
 XX  
 PT Immunogenic composition comprising OspC antigens - for the  
 PT treatment of Lyme borreliosis in different, specific geographical  
 PT areas.  
 XX  
 XX Disclosure; Fig. 9a; 115pp; English.  
 PS  
 XX A vaccine for Lyme disease includes selected OspC antigen  
 CC formulations based on defined OspC families resolved by serovar  
 CC typing and RFLP typing. Partial sequences of OspC genes selected  
 CC from different RFLP types are given in AAQ73883-905 (encoded peptides,  
 CC comprising the first 92% of mature OspC, are given in AA62771-93).  
 CC Complete sequences of these novel ospC genes, including the 3' end,  
 CC plus sequences for the ospC genes of Borrelia strains H13 and 28691  
 CC are given in AAQ73857-82, and encoded proteins in AA60884-909. The  
 CC DNA sequences may be expressed in e.g. Pichia pastoris for  
 CC recombinant antigen production.  
 XX  
 SQ Sequence 192 AA;

Query Match 3.2%; Score 7; DB 15; Length 192;  
 Best Local Similarity 100.0%; Pred. No. 72;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 39 AKEAILK 45  
 DB 137 AKEAILK 143

RESULT 19  
 AAB62703  
 ID AAB62703 standard; Protein; 192 AA.

XX AAB62703;  
 AC  
 XX 03-APR-2001 (first entry)  
 DT  
 XX B burgdorferi ospC protein SEQ ID NO: 10.  
 DE  
 XX Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.  
 KM  
 XX Borrelia burgdorferi.  
 OS  
 XX WO200078966-A1.  
 PN  
 XX 28-DEC-2000.  
 PD  
 XX 19-JUN-2000; 2000MO-US16915.  
 PF  
 XX 18-JUN-1999; 99US-0140042.  
 PR  
 XX (UNIV ) UNIV NEW YORK STATE RES FOUND.  
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.  
 XX  
 PI Datwyler RJ, Seinst G, Dykhuizen D, Luft BJ, Gomes-Solecki M;  
 XX WPI; 2001-050113/06.  
 DR N-PSDB; AAF29007.

XX Compositions of OspC polypeptides from strains of Borrelia which cause  
 PT Lyme disease are used to immunize animals and detect immune responses  
 PT to Lyme disease -  
 PT  
 XX  
 XX Disclosure; Page 67; 160pp; English.  
 PS  
 XX The present invention provides compositions comprising ospC proteins and  
 CC chimeric ospC proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads  
 CC to Lyme disease.  
 XX  
 SQ Sequence 192 AA;

Query Match 3.2%; Score 7; DB 22; Length 192;  
 Best Local Similarity 100.0%; Pred. No. 72;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 39 AKEAILK 45  
 DB 144 AKEAILK 150

RESULT 20  
 AAB62721  
 ID AAB62721 standard; Protein; 192 AA.

XX AAB62721;  
 AC  
 XX 03-APR-2001 (first entry)  
 DT  
 XX B burgdorferi ospC protein SEQ ID NO: 46.  
 DE  
 XX  
 XX Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.  
 KM  
 XX Borrelia burgdorferi.  
 OS  
 XX WO200078966-A1.  
 PN  
 XX 28-DEC-2000.  
 PD  
 XX 19-JUN-2000; 2000MO-US16915.  
 PF  
 XX 18-JUN-1999; 99US-0140042.  
 PR  
 XX (UNIV ) UNIV NEW YORK STATE RES FOUND.  
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.  
 XX  
 PI Datwyler RJ, Seinst G, Dykhuizen D, Luft BJ, Gomes-Solecki M;  
 XX WPI; 2001-050113/06.  
 DR N-PSDB; AAF29025.

XX Compositions of OspC polypeptides from strains of Borrelia which cause  
 PT Lyme disease are used to immunize animals and detect immune responses  
 PT to Lyme disease -  
 PT  
 XX  
 XX Disclosure; Page 105-106; 160pp; English.  
 PS  
 XX The present invention provides compositions comprising ospC proteins and  
 CC chimeric ospC proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads  
 CC to Lyme disease.  
 XX  
 SQ Sequence 192 AA;

Query Match 3.2%; Score 7; DB 22; Length 192;  
 Best Local Similarity 100.0%; Pred. No. 72;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 39 AKEAILK 45

Db |||||  
138 AKEALRK 144

## RESULT 21

AAG28190  
ID AAG28190 standard; Protein; 200 AA.

XX AAG28190;

DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 33316.

DE Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 29-MAR-1999; 99US-0126264.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 16-APR-1999; 99US-0128874.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 18-MAY-1999; 99US-0134370.

PR 19-MAY-1999; 99US-0134768.

PR 20-MAY-1999; 99US-0134941.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136332.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 14-JUN-1999; 99US-0138847.

PR 16-JUN-1999; 99US-0139452.

PR 17-JUN-1999; 99US-0139453.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.



```

PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149923.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0158295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match 3.2%; Score 7; DB 21; Length 200;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 175 VLVDGLQ 181
Db 163 VLVDGLQ 169

```

```

RESULT 22
AAU57004
ID AAU57004 standard; Protein; 202 AA.

```

```

XX AC AAU57004;
XX AC 27-FEB-2002 (first entry)
XX DT
XX DE Propionibacterium acnes immunogenic protein #17900.
XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
XX KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX KW dermatological; osteopathic; neuroprotectant.
XX OS Propionibacterium acnes.
XX PN WO200181581-A2.
XX XX 01-NOV-2001.
XX PD
XX PF 20-APR-2001; 2001WO-US12865.
XX XX
XX PR 21-APR-2000; 2000US-199047P.
XX PR 02-JUN-2000; 2000US-208841P.
XX PR 07-JUL-2000; 2000US-216747P.
XX PA (CORI-) CORIXA CORP.
XX PI Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX PI L'maisemeuve J, Zhang Y, Jen S, Carter D;
XX DR WPI; 2001-616774/71.
XX DR N-PSDB; AAS59579.
XX PT Propionibacterium acnes polypeptides and nucleic acids useful for
XX PT vaccinating against and diagnosing infections, especially useful for
XX PT treating acne vulgaris -
XX PS Example 1; SEQ ID NO 18199; 1069pp; English.
XX CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
XX CC polypeptides. The proteins and their associated DNA sequences are used in
XX CC the treatment, prevention and diagnosis of medical conditions caused by
XX CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
XX CC P. acnes is also involved in infections of bone, joints and the central
XX CC nervous system, however it is particularly involved in the inflammatory
XX CC lesions associated with acne vulgaris. A method for detecting the
XX CC presence or absence of P. acnes in a patient comprises contacting a
XX CC sample with a binding agent that binds to the proteins of the invention
XX CC and determining the amount of bound protein in the sample. The
XX CC polypeptides may be used as antigens in the production of antibodies
XX CC specific for P. acnes proteins. These antibodies can be used to
XX CC downregulate expression and activity of P. acnes polypeptides and
XX CC therefore treat P. acnes infections. The antibodies may also be used as
XX CC diagnostic agents for determining P. acnes presence, for example, by
XX CC enzyme linked immunosorbent assay (ELISA).
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 202 AA;

```

```

Query Match 3.2%; Score 7; DB 22; Length 202;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 178 DGIQLIQ 184
Db 187 DGIQLIQ 193

```

```

RESULT 23
ABP30264
ID ABP30264 standard; Protein; 205 AA.

```

```

XX ABP30264;
AC
XX
XX 02-JUL-2002 (first entry)
XX
XX Streptococcus polypeptide SEQ ID NO 9704.
DE
XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
XX group A streptococcus; Streptococcus pyogenes; antibacterial;
XX antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS Streptococcus agalactiae.
XX
XX MO200234771-A2.
XX
XX 02-MAY-2002.
XX
XX 29-OCT-2001; 2001WO-GB04789.
XX
XX 27-OCT-2000; 2000GB-0026333.
XX 24-NOV-2000; 2000GB-0028727.
XX 07-MAR-2001; 2001GB-0005640.
XX
XX (CHIR-) CHIRON SPA.
XX (GENO-) INST GENOMIC RES.
XX
XX Telford J, Massignani V, Margarit Ros YI, Grandi G, Fraser C;
XX Tettelin H;
XX WPI; 2002-352536/38.
XX N-PSDB; ABN70895.
XX
XX New Streptococcus protein for the treatment or prevention of infection
XX or disease caused by Streptococcus bacteria, such as meningitis, and
XX for detecting a compound that binds to the protein.
XX
XX Claim 1; Page 4097; 4525pp; English.
XX
XX The invention relates to a protein (ABP25413-ABP30895) from group B
XX streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
XX (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
XX the specification. The proteins have antibacterial and antiinflammatory
XX activity. (1), nucleic acids encoding (1), ABN66044-ABN71526 and
XX antibodies that bind (1) are used in the manufacture of medicaments for
XX the treatment or prevention of infection or disease caused by
XX Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
XX Nucleic acids encoding (1) are used to detect Streptococcus in a
XX biological sample. (1) is used to determine whether a compound binds to
XX (1). A composition comprising (1) or a nucleic acid encoding (1), may be
XX used as a vaccine or diagnostic composition. The disease caused by
XX Streptococcus that is prevented or treated may be meningitis. Nucleic
XX acid encoding (1) may be used to recombinantly produce (1) and may be
XX used in gene therapy. Antibodies to (1) are used for affinity
XX chromatography, immunoassays, and distinguishing/identifying
XX Streptococcus proteins.
XX
XX Sequence 205 AA;
XX
XX Query Match 3.2%; Score 7; DB 23; Length 205;
XX Best Local Similarity 100.0%; Pred. No. 77;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 37 ELAKEAI 43
XX |||||
XX 74 ELAKEAI 80
XX
XX RESULT 24
XX AAR75728 standard; Protein; 209 AA.
XX
XX AAR75728;
XX

```

```

DT 31-JUL-1996 (first entry)
XX
XX B. burgdorferi strain K48 outer surface protein C (OspC-K48).
DE
XX Strain K48; outer surface protein; OspC; antigenic domain;
XX chimeric protein; treatment; diagnosis; infection; vaccine;
XX Lyme borreliosis; immunodiagnostic assay; antibody;
XX T-cell reactivity; chimeric.
XX
XX Borrelia burgdorferi.
XX
XX WO9512676-A1.
XX
XX 11-MAY-1995.
XX
XX 27-OCT-1994; 94WO-US12352.
XX
XX 29-APR-1994; 94US-0235836.
XX 01-NOV-1993; 93US-0148191.
XX
XX (ASU-) ASSOC UNIVERSITIES INC.
XX
XX Dunn JJ, Luft BJ;
XX WPI; 1995-215034/28.
XX N-PSDB; AAQ90715.
XX
XX Chimeric protein comprising 2 or more antigenic Borrelia
XX polypeptide(s) - useful in a vaccine against Lyme borreliosis and in
XX immuno-diagnostic assays
XX
XX Example 1; Fig 13; 200pp; English.
XX
XX The present sequence is the B. burgdorferi strain K48, outer
XX surface protein C (OspC-K48). Using chemical or enzymatic methods,
XX peptide fragments of OspC-K48 were prepd., and analysed by western
XX blot to assess their ability to bind different anti-OspC monoclonal
XX antibodies. The information obtd. was used to locate antigenic
XX domains in OspC-K48, the epitopes of which were mapped with the
XX aid of site directed mutagenesis. Identical analyses were performed
XX on a selection of Osp purified from a variety of B. burgdorferi
XX strains, the results from which were utilised in the prepn. of a
XX pool of antigenic Borrelia polypeptides, and corresponding
XX polynucleotides. Chimeric proteins comprising 2 or more antigenic
XX Borrelia polypeptides, that do not naturally occur in the same
XX protein, can be used in the treatment and diagnosis of Borrelia
XX infections, i.e. as a vaccine against Lyme borreliosis, in
XX immunodiagnostic assays to detect anti-Borrelia antibodies or to
XX measure T-cell reactivity.
XX
XX Sequence 209 AA;
XX
XX Query Match 3.2%; Score 7; DB 16; Length 209;
XX Best Local Similarity 100.0%; Pred. No. 78;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 39 AKEAILK 45
XX |||||
XX 154 AKEAILK 160
XX
XX RESULT 25
XX AAB62720 standard; Protein; 209 AA.
XX
XX AAB62720;
XX
XX 03-APR-2001 (first entry)
XX
XX B burgdorferi ospC protein SEQ ID NO: 44.
XX
XX Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX

```

OS Borrelia burgdorferi.  
 XX  
 XX WO200078966-A1.  
 XX  
 PD 28-DEC-2000.  
 XX  
 XX 19-JUN-2000; 2000WO-US16915.  
 XX  
 XX 18-JUN-1999; 99US-0140042.  
 XX  
 PA (UNIV ) UNIV NEW YORK STATE RES FOUND.  
 PA (BROO-) BROOK BIOTECHNOLOGIES INC.  
 XX  
 PI Dactwyler RJ, Seinst G, Dykhuizen D, Luft BJ, Gomes-Solecki M;  
 DR WPI; 2001-050113/06.  
 DR N-PsDB; AAF29024.  
 XX  
 PT Compositions of OspC polypeptides from strains of Borrelia which cause  
 PT Lyme disease are used to immunize animals and detect immune responses  
 PT to Lyme disease -  
 XX  
 PS Disclosure; Page 104; 160pp; English.  
 XX  
 CC The present invention provides compositions comprising ospC proteins and  
 CC chimeric ospC proteins from members of the Borrelia genus. These may be  
 CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as  
 CC vaccines against Borrelia infection, which is spread by ticks and leads  
 CC to Lyme disease.  
 XX  
 SQ Sequence 209 AA;  
 Query Match 3.2%; Score 7; DB 22; Length 209;  
 Best Local Similarity 100.0%; Pred. No. 78;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 39 AKAALIK 45  
 DB 154 AKAALIK 160  
 RESULT 26  
 ID AAW11935  
 AC AAW11935 standard; Protein; 210 AA.  
 AC AAW11935;  
 DT 01-APR-1997 (first entry)  
 DE Outer surface protein C #2.  
 DE  
 KW Outer surface protein; ospC; Borrelia burgdorferi; IGM reactive; p23;  
 KW tick; spirochete; Lyme disease; multisystem infection; Lyme borreliosis;  
 KW immunodominant antigen; humoral antibody response; mammal; vaccine;  
 KW therapy.  
 OS Borrelia burgdorferi strain B31.  
 OS  
 PN WO9514781-A2.  
 PD 01-JUN-1995.  
 PF 22-NOV-1994; 94WO-US13540.  
 PF  
 PR 24-NOV-1993; 93US-0158353.  
 PR  
 PA (UYCO-) UNIV CONNECTICUT.  
 PI Padula SJ;  
 PI WPI; 1995-206938/27.  
 DR New Borrelia burgdorferi DNA encodes the outer surface protein C -  
 PT

PT used for early diagnosis and prevention of Lyme disease  
 XX  
 PS Claim 18; Page 44-45; 57pp; English.  
 XX  
 CC This sequence represents the Borrelia burgdorferi strain B31 outer  
 CC surface protein C (OspC). OspC is a 23 kD IGM reactive protein, and  
 CC is also known as p23. B31 is a North American B. burgdorferi strain,  
 CC that produces large amounts of OspC. B. burgdorferi is a tick borne  
 CC spirochete that causes Lyme disease (which is a multisystem infection).  
 CC OspC is an immunodominant antigen that elicits a humoral antibody  
 CC response in mammals with early Lyme disease. This OspC protein (in  
 CC recombinant or pure form) can be used for the specific and sensitive  
 CC detection of early Lyme disease. This is due to the fact that OspC is an  
 CC immunodominant antigen. The protein can also be used in a vaccine for  
 CC preventing Lyme borreliosis in mammals. Antibodies against this sequence  
 CC can be used to detect B. burgdorferi.  
 XX  
 SQ Sequence 210 AA;  
 Query Match 3.2%; Score 7; DB 16; Length 210;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 39 AKAALIK 45  
 DB 155 AKAALIK 161  
 RESULT 27  
 ID AAR75727  
 AC AAR75727 standard; Protein; 210 AA.  
 AC AAR75727;  
 DT 31-JUN-1996 (first entry)  
 DE B. burgdorferi strain B31 outer surface protein C (OspC-B31).  
 DE  
 KW Strain B31; outer surface protein; OspC; antigenic domain;  
 KW chimeric protein; treatment; diagnosis; infection; vaccine;  
 KW Lyme borreliosis; immunodiagnostic assay; antibody;  
 KW T-cell reactivity; chimeric.  
 XX  
 OS Borrelia burgdorferi.  
 OS  
 FH Key Location/Qualifiers  
 FH CDS 1..633  
 FT /\*tag= a  
 FT  
 PN WO9512676-A1.  
 PD 11-MAY-1995.  
 PD  
 PF 27-OCT-1994; 94WO-US12352.  
 PF  
 PR 29-APR-1994; 94US-0235836.  
 PR 01-NOV-1993; 93US-0148191.  
 PR  
 PA (ASUT-) ASSOC UNIVERSITIES INC.  
 PI Dunn JJ, Luft BJ;  
 PI WPI; 1995-215034/28.  
 DR N-PsDB; AAQ90714.  
 XX  
 PT Chimeric protein comprising 2 or more antigenic Borrelia  
 PT polypeptide(s) - useful in a vaccine against Lyme borreliosis and in  
 PT immuno:diagnostic assays  
 XX  
 PS Example 1; Fig 12; 200pp; English.  
 XX  
 CC The present sequence is the B. burgdorferi strain B31, outer  
 CC surface protein C (OspC-B31). Using chemical or enzymatic methods,  
 CC

CC peptide fragments of OspC-B31 were prep'd, and analysed by western  
CC blot to assess their ability to bind different anti-OspC monoclonal  
CC antibodies. The information obt'd. was used to locate antigenic  
CC domains in OspC-B31, the epitopes of which were mapped with the  
CC aid of site directed mutagenesis. Identical analyses were performed  
CC on a selection of Osp purified from a variety of B. burgdorferi  
CC strains, the results from which were utilised in the prep'n. of a  
CC pool of antigenic Borrelia polypeptides, and corresponding  
CC polynucleotides. Chimeric proteins comprising 2 or more antigenic  
CC Borrelia polypeptides, that do not naturally occur in the same  
CC protein, can be used in the treatment and diagnosis of Borrelia  
CC infections, i.e. as a vaccine against Lyme borreliosis, in  
CC immunodiagnostic assays to detect anti-Borrelia antibodies or to  
CC measure T-cell reactivity.

SO Sequence 210 AA;

Query Match 3.2%; Score 7; DB 16; Length 210;  
Best Local Similarity 100.0%; Pred. No. 79;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 39 AKBALX 45  
Db 155 AKBALX.161

RESULT 28  
AAG28189  
ID AAG28189 standard; Protein: 228 AA.

XX AAG28189;  
AC  
XX 17-OCT-2000 (first entry)  
DT  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 33315.  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PP 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132487.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144335.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 23-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 26-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.

```

PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0148368.
PR 17-AUG-1999; 99US-0148175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.

```

```

PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161992.
PR 29-OCT-1999; 99US-0162142.

Query Match 3.2%; Score 7; DB 21; Length 228;
Best local similarity 100.0%; Pred. No. 85;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 175 VLVDGLQ 181
    |||||
Db 191 VLVDGLQ 197

RESULT 29
AAV07075
ID AAV07075 standard; Protein; 233 AA.
XX
AC AAV07075;
XX
DT 02-JUL-1999 (first entry)
XX
DE Renal cancer associated antigen precursor sequence.
XX
KW Cancer associated antigen; diagnosis; research; treatment; human;
KW breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
KW prostate cancer.
XX
OS Homo sapiens.
XX
PN W09904265-A2.
XX
PD 28-JAN-1999.
XX
PF 15-JUL-1998; 98WO-US14679.
XX
PR 22-JUN-1998; 98US-0102322.
PR 17-JUL-1997; 97US-0896164.
PR 10-OCT-1997; 97US-0061599.
PR 10-OCT-1997; 97US-0061765.
PR 10-OCT-1997; 97US-0948705.
PR 11-OCT-1997; 97GB-0021697.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
XX Chen Y, Gout I, Gure A, O'Hare M, Obata Y, Old LJ;
PI Pfrendrich M, Sahin U, Scanlan MJ, Stockert B;
PI Tureci O;
XX
XX WPI, 1999-132448/11.
XX
XX New isolated cancer associated nucleic acids and polypeptides -
PT isolated using sera from cancer patients, used to develop products
PT for the diagnosis, monitoring or treatment of cancers
PT
XX
XX Disclosure; Page 479; 787pp; English.
XX
XX The invention relates to a method for diagnosing a disorder characterised
XX by expression of a human cancer associated antigen precursor coded for by
XX a nucleic acid molecule (NAM). The method comprises: (a) contacting a
XX biological sample isolated from a subject with an agent that specifically
XX binds to the NAM, an expression product or a fragment of an expression
XX product complexed with an HLA molecule; and (b) determining the
XX interaction between the agent and the NAM or the expression product as a
XX determination of the disorder. The products and methods can be used in
XX the diagnosis, monitoring, research, or treatment of conditions
XX characterised by the expression of various cancer associated antigens.
XX The invention provides nucleic acid sequences and encoded polypeptides
CC

```

CC which are cancer associated antigen precursors expressed in human breast  
CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and  
CC lung cancer.  
XX  
SQ Sequence 233 AA;  
Query Match 3.2%; Score 7; DB 20; Length 233;  
Best Local Similarity 100.0%; Pred. No. 87;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 170 LNSGSVL 176  
Db 15 LNSGSVL 21  
RESULT 30  
ABP26037  
ID ABP26037 standard; Protein: 236 AA.  
AC ABP26037;  
XX  
AC ABP26037;  
XX  
DT 02-JUL-2002 (first entry)  
XX  
DE Streptococcus polypeptide SEQ ID NO 1250.  
XX  
KM Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;  
KM group A streptococcus; Streptococcus pyogenes; antibacterial;  
KM antiinflammatory; infection; vaccine; meningitis; gene therapy.  
XX  
OS Streptococcus agalactiae.  
XX  
XX  
XX WO200234771-A2.  
XX  
XX 02-MAY-2002.  
XX  
XX 29-OCT-2001; 2001WO-GB04789.  
XX  
PF 27-OCT-2000; 2000GB-0026333.  
XX  
PR 24-NOV-2000; 2000GB-0028727.  
XX  
PR 07-MAR-2001; 2001GB-0005640.  
XX  
XX  
PA (CHIR-) CHIRON SPA.  
PA (GENO-) INST GENOMIC RES.  
XX  
XX Telford J, Maignani V, Margarit Ros YI, Grandi G, Fraser C;  
PI Tectelin H;  
XX  
XX WPI; 2002-352536/38.  
XX  
XX N-PSDB; ABN66668.  
XX  
XX  
XX New Streptococcus protein for the treatment or prevention of infection  
PT or disease caused by Streptococcus bacteria, such as meningitis, and  
PT for detecting a compound that binds to the protein -  
XX  
XX  
XX  
XX Claim 1; Page 3281; 4525pp; English.  
XX  
XX The invention relates to a protein (ABP25413-ABP30895) from group B  
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS  
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
CC the specification. The proteins have antibacterial and antiinflammatory  
CC activity. (1), nucleic acids encoding (1), ABN66044-ABN71526 and  
CC antibodies that bind (1) are used in the manufacture of medicaments for  
CC the treatment or prevention of infection or disease caused by  
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
CC Nucleic acids encoding (1) are used to detect Streptococcus in a  
CC biological sample. (1) is used to determine whether a compound binds to  
CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be  
CC used as a vaccine or diagnostic composition. The disease caused by  
CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
CC acid encoding (1) may be used to recombinantly produce (1) and may be  
CC used in gene therapy. Antibodies to (1) are used for affinity  
CC chromatography, immunoassays, and distinguishing/identifying  
CC Streptococcus proteins.

XX  
SQ Sequence 236 AA;  
Query Match 3.2%; Score 7; DB 23; Length 236;  
Best Local Similarity 100.0%; Pred. No. 88;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 37 ELAKEAI 43  
Db 105 ELAKEAI 111  
RESULT 31  
AAV29487  
ID AAV29487 standard; Protein: 238 AA.  
XX  
AC AAV29487;  
XX  
XX  
XX 13-OCT-1999 (first entry)  
XX  
XX Human lung tumour protein LTR6-6 predicted amino acid sequence.  
XX  
XX Human lung tumour protein; therapy; diagnosis; lung cancer; vaccine;  
XX immunotherapy; detection; inhibition.  
XX  
XX Homo sapiens.  
XX  
XX WO9938973-A2.  
XX  
XX  
XX 05-AUG-1999.  
XX  
XX 26-JAN-1999; 99WO-US01642.  
XX  
XX 22-DEC-1998; 98US-0219245.  
XX  
PR 28-JAN-1998; 98US-0015022.  
XX  
PR 28-JAN-1998; 98US-0015029.  
XX  
PR 18-MAR-1998; 98US-0040828.  
XX  
PR 18-MAR-1998; 98US-0040831.  
XX  
PR 23-JUL-1998; 98US-0122191.  
XX  
PR 23-JUL-1998; 98US-0122192.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX  
XX Frudakis TN, Lodes MJ, Mohamath R, Reed SG;  
PI  
XX WPI; 1999-479187/40.  
XX  
XX N-PSDB; AA207156.  
XX  
XX  
XX Lung tumour specific polynucleotides for inhibiting the development  
PT of lung cancer  
XX  
XX  
XX Example 2; Page 73-74; 171pp; English.  
XX  
XX The present invention describes lung tumour specific polynucleotides  
CC and tumour antigens. AA207144 to AA207246 and AA208301 to AA208325  
CC represent specifically claimed polynucleotides, and AAV29486 to AAV29571  
CC represent amino acid sequences from the present invention. The lung  
CC tumour specific polynucleotides and polypeptides can be used in  
CC pharmaceutical compositions and vaccines to inhibit the development of  
CC lung cancer. They can also be used to detect lung cancer in a patient.  
CC Probes and antibodies derived from the lung tumour sequences are useful  
CC in detection of lung cancer.  
XX  
XX  
XX Sequence 238 AA;  
Query Match 3.2%; Score 7; DB 20; Length 238;  
Best Local Similarity 100.0%; Pred. No. 88;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 170 LNSGSVL 176  
Db 16 LNSGSVL 22

```

RESULT 32
AAB44412
ID AAB44412 standard; Protein; 238 AA.
XX
XX AAB44412;
AC
XX 05-FEB-2001 (first entry)
DT
XX Human lung tumour-specific antigen encoded by cDNA #22.
DE
XX Lung tumour protein; lung cancer; cytostatic; vaccine.
XX
XX Homo sapiens.
OS
XX WO200060077-A2.
PN
XX 12-OCT-2000.
PD
XX 30-MAR-2000; 2000WO-US08560.
PF
XX 02-APR-1999; 99US-0285323.
PR 09-AUG-1999; 99US-0370838.
PR 30-DEC-1999; 99US-0476235.
PR 03-MAR-2000; 2000US-0518809.
XX
XX (CORI-) CORIXA CORP.
PA
XX Reed SG, Lodes MJ, Mohamath R, Secrist H;
PI
XX WPI; 2000-638466/61.
DR N-PSDB; AAC79067.
XX
XX Novel lung tumor polypeptides and polynucleotides, useful for
PT detecting, monitoring or treating cancer, especially lung cancer -
XX
XX Example 2; Page 100; 243pp; English.
XX
XX The present sequence is given in a specification relating to compounds
CC for therapy and diagnosis of lung cancer. Polypeptides comprising at
CC least an immunogenic part of a lung tumour protein are disclosed.
CC The polypeptides are useful for inhibiting the development of cancer,
CC especially lung cancer. Samples of T cells expressing the polypeptides
CC may be used to inhibit the development of cancer. The polypeptides are
CC also useful for detecting and monitoring the progression of cancer,
CC especially lung cancer.
XX
XX Sequence 238 AA;
SQ
Query Match 3.2%; Score 7; DB 21; Length 238;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 170 LNSGSVL 176
DB 16 LNSGSVL 22

```

```

PN WO200172295-A2.
XX
XX 04-OCT-2001.
PD
XX 28-MAR-2001; 2001WO-US09991.
PF
XX 29-MAR-2000; 2000US-0538037.
PR 05-JUN-2000; 2000US-0588937.
PR 18-AUG-2000; 2000US-0640878.
PR 22-SEP-2000; 2000US-234517P.
PR 01-NOV-2000; 2000US-0704512.
PR 14-DEC-2000; 2000US-0738973.
XX
XX (CORI-) CORIXA CORP.
PA
XX Reed SG, Lodes MJ, Mohamath R, Secrist H, Benson DR, Indrias CY;
PI Henderson RA, Fling SP, Algate PA, Elliot M, Mannion J, Kalos MD;
XX
XX WPI; 2001-639201/73.
DR N-PSDB; AAD23142.
XX
XX New human lung-specific polynucleotides and polypeptides for the
PT diagnosis and treatment of disease e.g. lung cancer -
XX
XX Example 2; Page 156; 378pp; English.
XX
XX The invention relates to isolated lung tumour-specific proteins and
CC their corresponding cDNA molecules. Lung tumour-specific proteins and
CC their antigen-presenting cells are useful for stimulating and/or
CC expanding T cells specific for a tumour protein, and for inhibiting
CC the development of cancer. The invention also relates to a composition
CC useful for stimulating an immune response, and for treating cancer. The
CC lung tumour specific oligonucleotide is useful in gene therapy and for
CC diagnosis, detection and treatment of lung cancer. The present sequence
CC is human lung tumour-specific protein.
XX
XX Sequence 238 AA;
SQ
Query Match 3.2%; Score 7; DB 22; Length 238;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 170 LNSGSVL 176
DB 16 LNSGSVL 22

```

```

RESULT 34
AAV29489
ID AAV29489 standard; Protein; 243 AA.
XX
XX AAV29489;
AC
XX 13-OCT-1999 (first entry)
DT
XX Human lung tumour protein LT86-8 predicted amino acid sequence.
DE
XX Human lung tumour protein; therapy; diagnosis; lung cancer; vaccine;
XX immunotherapy; detection; inhibition.
XX
XX Homo sapiens.
OS
XX WO9938973-A2.
PN
XX 05-AUG-1999.
PD
XX 26-JAN-1999; 99WO-US01642.
PF
XX 22-DEC-1998; 98US-0219245.
PR 28-JAN-1998; 98US-0015022.
PR 28-JAN-1998; 98US-0015029.
PR 18-MAR-1998; 98US-0040828.
PR 18-MAR-1998; 98US-0040831.

```

```
PR 23-JUL-1998; 98US-0122191.
PR 23-JUL-1998; 98US-0122192.
XX
XX (CORI-) CORIXA CORP.
PA
PI Frudakis TN, Lodes MJ, Mohamath R, Reed SG;
XX WPI; 1999-479187/40.
DR N-PSDB; AAZ07158.
XX
PT Lung tumour specific polynucleotides for inhibiting the development
XX of lung cancer
XX
XX Example 2; Page 75-76; 171pp; English.
XX
XX The present invention describes lung tumour specific polynucleotides
XX and tumour antigens; AAZ07144 to AAZ07246 and AAZ08301 to AAZ08325
XX represent specifically claimed polynucleotides, and AA129486 to AA129571
XX represent amino acid sequences from the present invention. The lung
XX tumour specific polynucleotides and polypeptides can be used in
XX pharmaceutical compositions and vaccines to inhibit the development of
XX lung cancer. They can also be used to detect lung cancer in a patient.
XX Probes and antibodies derived from the lung tumour sequences are useful
XX in detection of lung cancer.
XX
XX Sequence 243 AA;
SQ
Query Match 3.2%; Score 7; DB 20; Length 243;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 170 LNSGSVL 176
Db 30 LNSGSVL 36
|||
|||
|||

RESULT 35
AA129491
ID AA129491 standard; Protein; 243 AA.
XX
XX AA129491;
AC
XX 13-OCT-1999 (first entry)
DT
XX
XX Human lung tumour protein LT86-11 predicted amino acid sequence.
DE
XX
XX Human; lung tumour protein; therapy; diagnosis; lung cancer; vaccine;
XX immunotherapy; detection; inhibition.
XX
XX Homo sapiens.
OS
XX
XX WO9938973-A2.
PN
XX
XX 05-AUG-1999.
PD
XX
XX 26-JAN-1999; 99WO-US01642.
PF
XX
XX 22-DEC-1998; 98US-0219245.
PR
XX 28-JAN-1998; 98US-0015022.
PR
XX 28-JAN-1998; 98US-0015029.
PR
XX 18-MAR-1998; 98US-0040828.
PR
XX 18-MAR-1998; 98US-0040831.
PR
XX 23-JUL-1998; 98US-0122191.
PR
XX 23-JUL-1998; 98US-0122192.
PA
XX (CORI-) CORIXA CORP.
XX
XX Frudakis TN, Lodes MJ, Mohamath R, Reed SG;
XX WPI; 1999-479187/40.
DR N-PSDB; AAZ07160.
XX
XX Lung tumour specific polynucleotides for inhibiting the development
PT
```

```
PT of lung cancer
XX
XX Example 2; Page 78-79; 171pp; English.
XX
XX The present invention describes lung tumour specific polynucleotides
XX and tumour antigens; AAZ07144 to AAZ07246 and AAZ08301 to AAZ08325
XX represent specifically claimed polynucleotides, and AA129486 to AA129571
XX represent amino acid sequences from the present invention. The lung
XX tumour specific polynucleotides and polypeptides can be used in
XX pharmaceutical compositions and vaccines to inhibit the development of
XX lung cancer. They can also be used to detect lung cancer in a patient.
XX Probes and antibodies derived from the lung tumour sequences are useful
XX in detection of lung cancer.
XX
XX Sequence 243 AA;
SQ
Query Match 3.2%; Score 7; DB 20; Length 243;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 170 LNSGSVL 176
Db 21 LNSGSVL 27
|||
|||
|||

RESULT 36
AAB44414
ID AAB44414 standard; Protein; 243 AA.
XX
XX AAB44414;
AC
XX 05-FEB-2001 (first entry)
DT
XX
XX Human lung tumour-specific antigen encoded by cDNA #24.
DE
XX
XX Lung tumour protein; lung cancer; cytostatic; vaccine.
XX
XX Homo sapiens.
OS
XX
XX WO200060077-A2.
PN
XX
XX 12-OCT-2000.
PD
XX
XX 30-MAR-2000; 2000WO-US08560.
PF
XX
XX 02-APR-1999; 99US-0285323.
PR
XX 09-AUG-1999; 99US-0370838.
PR
XX 30-DEC-1999; 99US-0476235.
PR
XX 03-MAR-2000; 2000US-0518809.
XX
XX (CORI-) CORIXA CORP.
PA
XX
XX Reed SG, Lodes MJ, Mohamath R, Secrist H;
XX WPI; 2000-638466/61.
DR N-PSDB; AAC79069.
XX
XX Novel lung tumor polypeptides and polynucleotides, useful for
XX detecting, monitoring or treating cancer, especially lung cancer -
XX
XX Example 2; Page 101; 243pp; English.
XX
XX The present sequence is given in a specification relating to compounds
XX for therapy and diagnosis of lung cancer. Polypeptides comprising at
XX least an immunogenic part of a lung tumour protein are disclosed.
XX The polypeptides are useful for inhibiting the development of cancer,
XX especially lung cancer. Samples of T cells expressing the polypeptides
XX may be used to inhibit the development of cancer. The polypeptides are
XX also useful for detecting and monitoring the progression of cancer,
XX especially lung cancer.
XX
XX Sequence 243 AA;
SQ
```





XX 28-MAR-2001; 2001WO-US09991.  
 XX  
 PR 29-MAR-2000; 2000US-0538037.  
 PR 05-JUN-2000; 2000US-0588937.  
 PR 18-AUG-2000; 2000US-0640878.  
 PR 22-SEP-2000; 2000US-234517P.  
 PR 01-NOV-2000; 2000US-0704512.  
 PR 14-DEC-2000; 2000US-0738973.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Reed SG, Lodes MJ, Mohamath R, Secrist H, Benson DR, Indirias CY,  
 PI Henderson RA, Fling SP, Algate PA, Elliot M, Mannion J, Kalos MD;  
 XX  
 DR WPI; 2001-639201/73.  
 DR N-PSDB; AAD23147.  
 XX  
 PT New human lung-specific polynucleotides and polypeptides for the  
 PT diagnosis and treatment of disease e.g. lung cancer -  
 XX  
 PS Example 2; Page 159; 378pp; English.  
 XX  
 CC The invention relates to isolated lung tumour-specific proteins and  
 CC their corresponding cDNA molecules. Lung tumour-specific proteins and  
 CC their antigen-presenting cells are useful for stimulating and/or  
 CC expanding T cells specific for a tumour protein, and for inhibiting  
 CC the development of cancer. The invention also relates to a composition  
 CC useful for stimulating an immune response, and for treating cancer. The  
 CC lung tumour specific oligonucleotide is useful in gene therapy and for  
 CC diagnosis, detection and treatment of lung cancer. The present sequence  
 CC is human lung tumour-specific protein.  
 XX  
 SQ Sequence 243 AA;  
 Query Match 3.2%; Score 7; DB 22; Length 243;  
 Best Local Similarity 100.0%; Pred. No. 90;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 170 LNSGSVL 176  
 |||||  
 DB 21 LNSGSVL 27  
 RESULT 40  
 AAY29495  
 ID AAY29495 standard; Protein; 244 AA.  
 XX  
 AC AAY29495;  
 XX  
 DT 13-OCT-1999 (first entry)  
 XX  
 DE Human lung tumour protein LT86-15 predicted amino acid sequence.  
 XX  
 KW Human; lung tumour protein; therapy; diagnosis; lung cancer; vaccine;  
 KW immunotherapy; detection; inhibition.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W09938973-A2.  
 XX  
 PD 05-AUG-1999.  
 XX  
 PF 26-JAN-1999; 99WO-US01642.  
 XX  
 PR 22-DEC-1998; 98US-0219245.  
 PR 28-JAN-1998; 98US-0015022.  
 PR 28-JAN-1998; 98US-0015029.  
 PR 18-MAR-1998; 98US-0040828.  
 PR 18-MAR-1998; 98US-0040831.  
 PR 23-JUL-1998; 98US-0122191.  
 PR 23-JUL-1998; 98US-0122192.  
 XX

PA (CORI-) CORIXA CORP.  
 XX  
 PI Fridakis TN, Lodes MJ, Mohamath R, Reed SG;  
 XX  
 DR WPI; 1999-479187/40.  
 DR N-PSDB; AAZ07164.  
 XX  
 PT Lung tumour specific polynucleotides for inhibiting the development  
 PT of lung cancer  
 XX  
 PS Example 2; Page 81-82; 171pp; English.  
 XX  
 CC The present invention describes lung tumour specific polynucleotides  
 CC and tumour antigens. AAZ07144 to AAZ07246 and AAZ08301 to AAZ08325  
 CC represent specifically claimed polynucleotides, and AAY29486 to AAY29571  
 CC represent amino acid sequences from the present invention. The lung  
 CC tumour specific polynucleotides and polypeptides can be used in  
 CC pharmaceutical compositions and vaccines to inhibit the development of  
 CC lung cancer. They can also be used to detect lung cancer in a patient.  
 CC Probes and antibodies derived from the lung tumour sequences are useful  
 CC in detection of lung cancer.  
 XX  
 SQ Sequence 244 AA;  
 Query Match 3.2%; Score 7; DB 20; Length 244;  
 Best Local Similarity 100.0%; Pred. No. 90;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 170 LNSGSVL 176  
 |||||  
 DB 22 LNSGSVL 28

Search completed: April 9, 2003, 14:43:02  
 Job time : 96 secs



GenCore version 5.1.4.p5 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 8, 2003, 04:54:34 ; Search time 1360 Seconds

(without alignments)  
2596.041 Million cell updates/sec

Title: US-10-034-500-2

Perfect score: 1119

Sequence: 1 AEVTASCTKREVSYNLYVDY.....QEQILVTEVVVLGNVAF 218

Scoring table: BIOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Delop 6.0 , Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xjh

-Q=/cgn2\_1/USPTO.spool/US10034500/runat\_02042003\_092745\_19713/app\_query.faeta\_1.391

-DB=EST -QMT=fastap -SUFFIX=ret -MINMATCH=0.1 -LOOFCU=0 -LOOPEXT=0

-UNITS=bits -START=1 -END=1 -MATRIX=blousum62 -TRANS=human40.cdi -LIST=45

-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40 -MODE=LOCAL

-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000

-USER=US10034500 @CGN 1.1 1456 @runat\_02042003\_092745\_19713 -NCPU=6 -ICPU=3

-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120

-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7

-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:

1: em\_eebta:\*

2: em\_eebta:\*

3: em\_eebta:\*

4: em\_eebta:\*

5: em\_eebta:\*

6: em\_eebta:\*

7: em\_eebta:\*

8: em\_eebta:\*

9: gb\_eebta:\*

10: gb\_eebta:\*

11: gb\_eebta:\*

12: gb\_eebta:\*

13: gb\_eebta:\*

14: gb\_eebta:\*

15: em\_eebta:\*

16: em\_eebta:\*

17: gb\_gsa:\*

18: em\_gsa:\*

19: em\_gsa:\*

20: em\_gsa:\*

21: em\_gsa:\*

22: em\_gsa:\*

23: em\_gsa:\*

24: em\_gsa:\*

25: em\_gsa:\*

26: em\_gsa:\*

27: em\_gsa:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	91	8.1	779	BH063089
2	90	8.0	797	BG518685
3	89.5	8.0	1859	AK019914
4	88	7.9	831	BG286204
5	87.5	7.8	641	BH245004
6	87.5	7.8	1047	B0879640
7	87	7.8	501	B1361347
8	86.5	7.7	608	B1010953
9	86.5	7.7	876	BF691597
10	86	7.7	841	AZ683111
11	86	7.7	985	CNS03082
12	85.5	7.6	1411	BG491140
13	85	7.6	917	AZ203552
14	84.5	7.6	713	BM246457
15	84.5	7.6	801	BQ518708
16	84.5	7.6	835	BQ792380
17	84	7.5	588	B1611208
18	84	7.5	1144	BM911554
19	83.5	7.5	625	BE372805
20	83.5	7.5	907	CNS06W1A
21	83	7.4	395	A1545929
22	83	7.4	886	BH132626
23	82.5	7.4	573	B1540501
24	82.5	7.4	611	AZ934444
25	82.5	7.4	1007	CNS06W1A
26	82	7.3	405	BG463740
27	82	7.3	524	BE596201
28	82	7.3	560	BE599098
29	82	7.3	609	BE585416
30	82	7.3	695	BH538458
31	82	7.3	987	BF288073
32	82	7.3	2933	AK008659
33	81.5	7.3	531	BE75248
34	81.5	7.3	867	BF304340
35	81.5	7.3	918	B0732001
36	81.5	7.3	1713	AY104471
37	81	7.2	601	BG447332
38	81	7.2	628	BH563226
39	81	7.2	764	BG850892
40	81	7.2	764	A1684487
41	81	7.2	1071	BF246267
42	80.5	7.2	554	BE016028
43	80.5	7.2	966	BQ718678
44	80	7.1	515	AV954108
45	80	7.1	590	AZ790944

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
BH063089/c	BH063089	RPCT-24-355L8.TV RPCT-24 Mus musculus genomic clone RPCT-24-355L8, DNA sequence.	BH063089	BH063089	GI:114877343	house mouse.	Eukaryote; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.	1 (bases 1 to 779)	Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akincet,B., Levins,M., Tsagaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E.,



Best Local Similarity:	21.67%	Mismatches:	82
Query Match:	8.04%	Indels:	20
DB:	12	Gaps:	6

bulk excision from Lambda FLC 1. Cloning sites, 5' end: SalI, 3' end: BamHI. Host: DH10B.

## FEATURES

## SOURCE

1. 1859  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="PANTOM:DB:530427N04"  
/db\_xref="MGD:MG1:1899542"  
/db\_xref="taxon:10090"  
/clone="5330427N04"  
/sex="male"  
/tissue\_type="pituitary gland"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
166\_>1859  
/note="data source: SPRR, source key: 000534, evidence: ISS homolog to BREAST CANCER SUPPRESSOR CANDIDATE 1 putative"  
/codon\_start=1  
/protein\_id="BAB31913.1"  
/db\_xref="GI:12860312"  
/translation="MEHICGLITSNKETVPLKNISVTLSINEFVAIVATINENEK  
VPLEATFVPMDEDSAVSFALVDGKIVALELODKMAHSEYERALSQHQAYLLE  
DDYSDVSCVGNLQPGAKAVTLRYVQELPLETGAIRYLPLINPGVQLSOSA  
NSCLNIQKPTVLEDTPTLNTTATITSOHGERVQNSCSPIQYITDDTSQVSL  
TEGHKEDVDVLLIYNEVSPSVAVEMGLMDKRSIMGSAVMSVSPPIPEVAS  
KACGEFVFLMBSSGMSDPMSTENNSQRIEAKETLLILKSLPMGCYNNYGGGS  
YKPPESVKTQDTMEDAVKVKVKKALKNALGTELLTLCPLNYKASSTIGHDLQFLV  
TDEVSDFSVIREVANSKKRCPSPGIGQASLSLKNLARVSATVAVITGKDM  
OTKALGSLKFAQCANDNISLSMDLPSGLSVMSLSPEDLIITFRGRLITVQLGLMP  
QVESTGACVLKHIILQGRSLENVTFSLQPKANDNFTIRHLAKSLIQKDFSGQSAS"

## CDS

BASE COUNT 499 a 424 c 454 g 482 t

## ORIGIN

## Alignment Scores:

Pred. No.: 3.3 Length: 1859  
Score: 89.50 Matches: 55  
Percent Similarity: 43.65% Conservative: 31  
Best Local Similarity: 27.92% Mismatches: 88  
Query Match: 8.00% Indels: 23  
DB: 11 Gaps: 11

US-10-034-500-2 (1-218) x AK019914 (1-1859)

QY 2 GUAUATThAlaseCystrLyArgValGluSerTyRantTytleuValaApyTyRset 21  
DB 979 GAAGTGAAGCCCTCA-----AAGGCGCTGTGAAGAAATTGTCTTCATGAGCCGTTCA 1032  
QY 22 GAlSerMet-----MetMetLySHleValAlaValArgGluProLySileGluLeuAla 39  
DB 1033 GGAAGTThGACTCCCCCATGACACAGAACAAATTCAGTACGATCGAGCGCTGCC 1092  
QY 40 LySGlAlaAlaIleuLySileAsnAlaAlaMetProLyMetSerTyR-----55  
DB 1093 AAGGAACCTGTGTGCTGTGAAGAGTTGTGCTTATTTAATATCTAT 1152  
QY 56 ---GInGlyIleuTyR---ThrPheAlaProTySerValIlelleleProGInGlySer 73  
DB 1153 GGAATTGACTTCTCAATGAGAAATCTTCCGAGAGTGTGAAGTCACTCGAGATACA 1212  
QY 74 TrpAsnSerCyValAlaGluCyAlaValaValAsnThrIleLySerAspLeuGluIlePhe 93  
DB 1213 ATGAGAGACCAAGTGAAGAA-----GTAGAGCTTTAAAGCACTTThGGGGGTACT 1266  
QY 94 GUAArgLeuThrProValGlyArgGlyIleLyMetHisGluThrValIleAsnGlnMet 113  
DB 1267 GAATCTTGACACCCCTCGACATTTACAAG-----GCATCTTCATTTCTGTGCAT 1320  
QY 114 ProProGlnAlaAlaValIleleuLeuThrAspGlyHisAsnMetLeuGlyMetAsnPro 133  
DB 1321 CCGCTACAG-----CTCTTGTCTTCAAGATGAGAGAGTGAAGTGAACATTTAGAGTGC 1374

QY 134 ValGluGluValLySeriTyRInThraSnProAsnValCysPheHisValaValSer 153  
DB 1375 ATTAAGAGAGTTTAAG-----TTAAACAGCAAGAAACAGATCTTTC-----TCT 1419

QY 154 PheAla---AspAspAlaGluGlyLyValAlaIlelleleAspGlnIleValAlaLeuAsnSer 172

DB 1420 TTTGGAAATTGACACAGAGCCCTGACACCGATTAAATTAATATGCGCGGATTCAGGG 1479

QY 173 GLY---SerValIleuValaAspGlyLeuGlnIleuLeuGlnAsnProAlaVal 188  
DB 1480 GGTACTGACGTCTTTATATACAGGACAGACAGATGACAGCAAGGCTCTT 1530

## RESULT 4

## LOCUS

BG286204 831 bp mRNA linear EST 21-FEB-2001

## DEFINITION

602383026F1 NIH\_MGC\_93 Homo sapiens cDNA clone IMAGE:4500439 5',

## ACCESSION

BG286204

## VERSION

BG286204.1 GI:13038883

## KEYWORDS

## SOURCE

## ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

1 (bases 1 to 831)  
NIH-MGC <http://mgs.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMNL at: <http://image.lnl.gov>  
Plate: LAM10366 row: d column: 08  
High quality sequence stop: 685.

## FEATURES

## SOURCE

1. 831  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4500439"  
/clone\_lib="NIH MGC 93"  
/tissue\_type="transitional cell papilloma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: bladder; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.7 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH\_MGC library."

## BASE COUNT

246 a 174 c 192 g 219 t

## ORIGIN

## Alignment Scores:

Pred. No.: 1.37 Length: 831  
Score: 88.00 Matches: 35  
Percent Similarity: 38.82% Conservative: 24  
Best Local Similarity: 23.03% Mismatches: 43  
Query Match: 7.86% Indels: 50  
DB: 12 Gaps: 5

US-10-034-500-2 (1-218) x BG286204 (1-831)

QY 43 ILeuLySileAsnAlaAlaMetProLyMetSerTyRInGInGlyIleuTyRThrPhe 62  
DB 142 CTATCAATTAATAAAGAGATGAGAAACACACTCATGCGAGATTAACCTCATAT 201  
QY 63 AlAProTySerValIlelleleProGInGlySeriTrpAsnSerCyValAlaGluCyAla 82  
DB 202 -----CCTCTGGAGAGAACTTCATCTGC----- 225

Qy 83 ValaAntHrIleuSerAspLeuGluIlePheGlyArgLeuThrProValGlyAspGly 102  
Db 226 ---TCGGAATTAAATATGATTCATTCAGAGATTTGAGAGCTACATTCCTCACTCGATGGA 282  
Qy 103 IleuSerMetHisGluThrValIleAsnGlnMetProProGlnAlaAlaValIleuLeu 122  
Db 283 -----TCGGAAGTACTGCTGCTG 300  
Qy 123 ThrAspGlyHisAsnAsnLeuGlyMetAsnProValGluGluValIleSerIleTyrgln 142  
Db 301 ACTGATGGGAGGATACACCTGCAAGTTCTTGTATTGATGAAGTAAA-----CA 351  
Qy 143 ThrAsnProAsnValCysPheHisValIleSerPheAla----- 155  
Db 352 AGTGGGGCATTTGTTCAATTTATGCTTTGGAGAGAGCTGTGATGAAGCACTAATAGAG 411  
Qy 156 -----AspAspAlaGluGlyVal 162  
Db 412 ATGACCAAGATACAGAGAGAGATCATTTTATGTTTCAGATGAAGCTCAAGAACATGGC 471  
Qy 163 IleIleAspGlnIleValAlaLeuAsnSerGlySer 174  
Db 472 CTCATTGATGCTTTGGGGCTCTTACATCAGAAAT 507

RESULT 5  
BH245004/c 641 bp DNA linear GSS 13-NOV-2001

LOCUS AUBA05TR AUBA05TR thaliana genomic clone AUBA05, DNA  
DEFINITION sequence.

ACCESSION BH245004  
VERSION BH245004.1 GI:16922378

KEYWORDS GSS.

SOURCE chae creas.  
ORGANISM Arabidopsis thaliana

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
1 (bases 1 to 641)

AUTHORS Town, C.D., Whitelaw, C.A., Pal, G., Van Aken, S.E., Uteback, T.V.,  
Feldblyum, T.V. and Fraser, C.M.

TITLE Survey sequencing of Arabidopsis thaliana BAC T7013  
JOURNAL Unpublished (2001)

COMMENT Other GSSs: AUBA05TR  
Contact: Chris Town

TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org

From Wash. U contig 1131.  
Seq primer: TP

Classes: sheared ends.

FEATURES location/Qualifiers

1..641

/organism="Arabidopsis thaliana"  
/strain="Columbia"  
/db\_xref="taxon:3702"

/clone="AUBA05"

/clone\_lib="AUBB"  
/note="Vector: pHOS2, Site\_1: BstXI, 2-3 kb sheared BAC  
DNA inserted into pHOS2 using BstXI linkers"

BASE COUNT 140 a 147 c 138 g 216 t

ORIGIN

Alignment Scores:

Pred. No.: 1.03 Length: 641

Score: 87.50 Matches: 44

Percent Similarity: 33.75% Conservatve: 37

Best Local Similarity: 18.33% Mismatches: 57

Query Match: 7.82% Indels: 102

DB: 17 Gaps: 9

US-10-034-500-2 (1-218) x BH245004 (1-641)

Qy 2 GluValThrAlaSerCysThrLysArgValGluSerTyrAsnTyr----- 16  
Db 617 GAATTAACAGCTGAATTCAGAGAAAGTCAATACATTCGCTGCTCCCGAAGATCTG 558  
Qy 17 -----LeuValAspTyrSerGlySerMetMetLysHisValAla 30  
Db 557 CCAGAGTAAAGTGTGATATGTCATCACAGCTCAACGACCTCAACTT 498  
Qy 31 ValArgGluProLysIleGluLeuAlaLysGluAlaIleLeuLysIleAsnAlaLamet 50  
Db 497 GTTAAGCAGAAAGAGAGAAATTTGGTCGGAGCAGACAGAACCGTAAACCGAAGTA 438  
Qy 51 ProLysMetSerTyrGlnGlyIleuTyThrPheAlaProTyr----- 65  
Db 437 GAGAAATTACAGAGATGGGTCATC---ACGAGCGAAATATTCGATTGACTGCG 381  
Qy 66 ---SerValIleIleProGln-----GlySerTyrPheAsnSerCysVal----- 78  
Db 380 AACTCAGTCGTATTAAAGAAAGAAACCGTAAATGAGAGCTGCTGATTCACCGAC 321  
Qy 79 -----AlaGluCysAlaVal--- 83  
Db 320 CTAAACAAAGCTTACCCAAAGACAGTTCCCATTAACCGACCTCGATTGCTGTAGAA 261  
Qy 84 -----AsnThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProVal 99  
Db 260 TCAACTTCTGAAACAGTTTGTTCATTCATGACGCTTTTGGC----- 216  
Qy 100 GlyAspGlyIleLysMetHisGluThrValIleAsnGlnMetProProGlnAlaAlaVal 119  
Db 216 ----- 216  
Qy 120 IleuLeuThrAspGlyHisAsnAsnLeuGlyMetAsnProValGluValIleSer 139  
Db 215 -----GGTTACACACAGATCATGAGAACCCCGGATCAAGAAACCC 171  
Qy 140 IleTyrglnThrAsnProAsnVal---CysPheHisValIleSerPheAlaAspAspAla 158  
Db 170 GCATTCTACACAGAAACAGCATCTTTGTTACCGAGTGAAGCCATTCCGA----- 120  
Qy 159 GluGlyLysAlaIleIleAspGlnIleValAlaLeuAsnSerGlySerValLeuValAsp 178  
Db 120 ----- 120  
Qy 179 GlyLeuGlnLeuLeuGlnAsnProAlaValCysGlnGluLysValAsnSerValPhe 197  
Db 119 -----CTCAAAACGCTGGCAACTTATCAACGCTTCACAAATCTTT 74

RESULT 6  
B0879640 1047 bp mRNA linear EST 16-AUG-2002

LOCUS B0879640  
DEFINITION AGENCOURT\_8119823 Lupski\_dorsal\_root\_ganglion Homo sapiens cDNA  
clone IMAGE:617787 5', mRNA sequence.

ACCESSION B0879640  
VERSION B0879640.1 GI:22271648

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1047)

AUTHORS NIH-MGC http://mgs.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@nci.nih.gov

CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:



http://image.llnl.gov  
 Plate: LLM13556 row: 1 column: 14  
 High quality sequence stop: 332.  
 Location/Qualifiers

# FEATURES

source

```
1..1047
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6177877"
/clone_lib="lupski_dorsal_root_ganglion"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/Note="Vector: PCMV-SPORT6 (Life Technologies); Site 1:
NotI; Site 2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCAGGCTCCG-3' and
5'-GACTAGTTCTAGATCGCGAGCGCCCTT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."
```

BASE COUNT 297 a 274 c 280 g 196 t  
 ORIGIN

## Alignment Scores:

Pred. No.:	2.33	Length:	1047
Score:	87.50	Matches:	45
Percent Similarity:	45.70%	Conservative:	40
Best Local Similarity:	24.19%	Mismatches:	59
Query Match:	7.82%	Indels:	42
DB:	14	Gaps:	9

US-10-034-500-2 (1-218) x BQ879640 (1-1047)

```

QY 17 LeuValAspTyrSerGlySerMetMetCysVal-----AlaVal 31
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 GTCCTGACCTTCAAGCTTCAAGCATCTACTGTGTAGATGATGACAGCATT 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 32 ArgGluProLysIleGluLeuAlaLysGluAlaIleLeuLysIleAsnAlaMetPro 51
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 121 GGGGCGACGAACTTCAAGAGGCAAAAGTCTAGTCAACTTA-----ATTGAG 171
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 52 LysMetSerTyrGlnGly-----GlyLeuTyrThrPheAlaProTyrSer 66
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 172 AAGGTGCAAGTATGATGTGAAGCAAGATATGCTAGTGCATATGCAATACCCC 231
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 67 ValIleIleProGlnLysSerTyrAsnSerCysValAlaGluCysAlaVal----- 83
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 232 AAAATT-----TGGGTCAAGTGTCTGAAGCAGACAGCATATGCAAGC 276
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 84 -----AsnThrIleLysSerAspLeuGlnIlePheGly-----ArgLe 96
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 277 TGGGTGACGAAAGCATCTCAATGAATCAATATGAAGACCAAGATTGAAGTCAGGACT 336
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 96 uThrProValGly---AspGlyIleLysMetHisGluThrValIleAsnGlnMetPro 114
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 337 AACACCAAGAAAGCCCTCCAGGAGTGTACAGATATGAGTGGCCAGATGACGTCCCT 396
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 115 ProGlnAla-----AlaValIleLeuLeuThrAspGlyHisAsnAsn 128
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 397 CCGAAGGCTGGAACCGACCCGCATGTATCTCTCTGACTGATGATGATGACACAC 456
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 129 LeuGlyMetAsnProValGluGluValLysSerIleTyrGlnThrAsnProAsnVal 148
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 457 ATGGGGGGGAGCCCAATTACTGTGATGATGATGATGATGATGATGATGATGATGAT 504
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 149 PheHisValValSerPheAlaAspAlaGluGlyLysAlaIleLeuAspGlnIleVal 168
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 505 TATATCATTTGGGGAAGGATCCCAAAACCCAGGGGAGGATTTATCCGGATGTCTAAGG 564
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 169 AlaLeuAsnSer 172
```

DB 565 GGCTTGGCGGCC 576

# FEATURES

source

```

RESULT 7
LOCUS B0361347 501 bp mRNA linear EST 07-MAR-2002
DEFINITION B0361347 Dictyostelium discoideum cDNA library, CF Dictyostelium
ACCESSION B0361347
VERSION B0361347.1 GI:19260942
KEYWORDS EST.
SOURCE Dictyostelium discoideum.
ORGANISM Dictyostelium discoideum.
REFERENCE Eukaryota; Mycetozoa; Dictyostelidae; Dictyostelium.
AUTHORS 1 (bases 1 to 501)
TITLE Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.
Full length cDNA of Dictyostelium discoideum at the culmination
stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasi Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yatae, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: teshin@genes.nig.ac.jp.
```

## Alignment Scores:

Pred. No.:	0.798	Length:	501
Score:	87.00	Matches:	39
Percent Similarity:	54.23%	Conservative:	38
Best Local Similarity:	27.46%	Mismatches:	46
Query Match:	7.77%	Indels:	19
DB:	13	Gaps:	9

US-10-034-500-2 (1-218) x B0361347 (1-501)

```

QY 80 GluCysAlaValAsnThrIleLysSerAspLeuGlnIlePheGly---ArgLeuThr--- 97
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 42 AAGTGTATTTAGTATTCATGTAGA---CATTTAGAGATGAGGAGATTAAAGTACTTAT 98
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 98 ---ProValGlyAsp---GlyIleLysMetHisGluThrValIleAsnGlnMetPro 115
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 99 TTACAGTGGGTGAAGAAAGTCTTACATTTAGGATTAGATTAGCAAGCAATTCACACA 158
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 116 GlnAlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGly---MetAsnProVal 134
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 159 CAACATCATTTGTGAACATTTATGCGAGTTATTAATTAATGAGATGGTGTGAACCTATT 218
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 135 GluGluValLysSerIleTyrGlnThrAsnProAsnValCysPheHisValValSerPhe 154
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 219 AAAGAGATTGTAATTTGTAGAGAGAT---GGGTGTTTCTTCAC-----ACCGAT 269
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 155 AlaAspAspAlaGluGlyLysAlaIleLeuAsp-----GlnIleVal 168
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 270 GCTGCAAGCAGATGTTGTAATGACCAATGATGATGATGATGATGATGATGATGATGAT 329
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 169 AlaLeuAsnSerGlySerVal-LeuValAspGlyLeuGlnLeuGlnAsnProAlaVal 188
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 330 TCGATTAGGGGTACAAAATATACGTTCCAAAGGGTGGGATTTATTCGTACAGACT 389
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 188 LysGlnGluPheValAsnSerValPheCysGlnGluGlnIleLeuValThrGluGluVal 208
```

Db	330	AGACCAAGTACGTATCGACG-----CAATCAACTGCTGCTGCTCAAGACT	440
Qy	208		
Db	441	GGTA 444	
RESULT 8			
LOCUS	BJ109553		
DEFINITION	BJ109553	608 bp	mRNA linear EST 23-JAN-2002
ACCESSION	BJ109553	unpublished oligo-capped cDNA library, C. elegans L1 stages	
VERSION	BJ109553	Caenorhabditis elegans cDNA clone yk1120f07 5', mRNA sequence.	
KEYWORDS	EST.		
SOURCE	BJ109553.1	GI:1826582	
ORGANISM	Caenorhabditis elegans.		
REFERENCE	Caenorhabditis elegans.		
AUTHORS	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae		
TITLE	Rhabditidae; Peloderinae; Caenorhabditis.		
JOURNAL	1 (bases 1 to 608)		
COMMENT	Kohara, Y., Shin-i, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y. and Sugano, S.		
FEATURES	A complementary view of the C. elegans genome		
source	Unpublished (2002)		
	Contact: Tadasi Shin-i		
	Center for Genetic Resource Information		
	National Institute of Genetics		
	1111 Yata, Mishima, Shizuoka 411-8540, Japan		
	Tel: 81-559-81-6856		
	Fax: 81-559-81-6855		
	Email: tehin@genes.nig.ac.jp.		
	Location/Qualifiers		
	1. .608		
	/organism="Caenorhabditis elegans"		
	/strain="N2"		
	/db_xref="taxon:6239"		
	/clone="yk1120f07"		
	/clone_lib="unpublished oligo-capped cDNA library, C.		
	elegans L1 stage"		
	/sex="hermaphrodite"		
	/tissue_type="whole animal"		
	/dev_stage="L1"		
	/note="The AD-wrmcDNA library was generated with poly(A)+		
	mRNA isolated from both hermaphrodite and male N2 worms of		
	all larval stages, embryos, adults and dauers and the		
	subsequent generation of cDNAs by poly(A) priming. The		
	cDNAs were cloned into pC86"		
BASE COUNT	208 a 131 c 120 g 149 t		
ORIGIN			
Alignment Scores:			
Pred. NO.:	1.28	Length:	608
Score:	86.50	Matches:	43
Percent Similarity:	37.86%	Conservative:	35
Best Local Similarity:	20.87%	Mismatch:	67
Query Match:	7.73%	Indels:	61
DB:	13	Gaps:	9
US-10-034-500-2 (1-218) x BJ109553 (1-608)			
Qy	34	ProlysvlllegluLeuAlaLyglu--AlaileuLyllleasAlaAlaMetProlys	52
Db	11	CCCAAGTTTGAGATGCTACAGAGAGTGCACATCAACAAAGTTACAGTAAATCATCCCGTG	70
Qy	53	MetserTYTGinglyGlyleuTythrPheAlaProTYrSer-----	65
Db	71	AACAATGATGACGTTGAGATGCTGCGGATGATCATGTTTCAGAACTAAAGACAAGTT	130
Qy	67	---VallelleProGinglyserTrpaenserCyvalAlaGlucyAlaValaenThr	85
Db	131	CTTGCTTCCTCCAACTCCGAACAAGACAAGTT-----TGTATATATTACACC	131
Qy	86	-----lleYserAspIeuGluIlePheGlyArgLeuThrProValGlyAspGlyIle	103

Db	182	GGAAAAATTCGAAAGGAAAGAAAGAAATCTGCACAG---CACAAATGCTGATGTGCAC	238
Qy	104	LYSMETHISGLUTHIValIleAeNGInMeProProGInaIaIaValIleLeuDeuThr	123
Db	239	ACCGTCACCTTGATATTATGAATATCAAGCCGGTCCAAACACAGGCGCGGCTGTCCAA	298
Qy	123	-----	123
Db	299	CCAAAGCTTCATCTGCACCAACAATGGAGAAATGGAGTACCTGTATATTAAAC	358
Qy	124	-----AspGLyHISaenAsnLeuGLyMeAsnProVal	134
Db	339	AATCCAGATGCTATGCCGTTCAGTTATGGATATATCAATATTCACAAACAATCTTAGAAAT	418
Qy	135	---GluGluValYLeSerIleTyrgInThxAnPProAnValCywPheIleValIser	153
Db	419	CSAGGTTATAGAGAACATATTCAATCCAAACCA-----	454
Qy	154	PheAlaAspAlaIleGluGLyValAlaIleIleAspGInIleVal-----Ala	169
Db	455	-----CAATTCCAAGCATGTGATTGGAGAAATCCAGAAATGGATGCACAT	499
Qy	170	LeuAsnSerGLySerValLeuValAspGLyLeuGInLeuGInAsnProAlaValCyS	189
Db	500	CTCAACGATCCGAATGATATGATCGTCAAACTATGAGATGATTCGTATTCAAATATGTTC	559
Qy	190	GInGluPheValAsnSer	195
Db	560	CAGGAATGATGCCGAAT	577
RESULT 9			
LOCUS	BF691597	876 bp	mRNA linear EST 22-DEC-2000
DEFINITION	602248190F01 NIH_MGC_62 Homo sapiens cDNA clone IMAGE:433330 5',		
ACCESSION	BF691597		
VERSION	BF691597.1	GI:11977005	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 876)		
JOURNAL	NH-MGC http://mgc.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Straubeberg, Ph.D.		
	Email: cgabbs-remail.nih.gov		
	Tissue Procurement: ATCC/DCTO/DRP		
	cDNA Library Preparation: CLONTECH Laboratories, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LMNL at:		
	http://image.lnl.gov		
	Plate: LCM1201 row: C column: 23		
	High quality sequence stop: 601.		
FEATURES	Location/Qualifiers		
SOURCE	1..876		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:433330"		
	/clone_1db="NIH MGC 62"		
	/tissue_type="melanotic melanoma, high MDR"		
	/lab_host="DH10B (T1 phage-resistant)"		
	/note="Organ: skin; Vector: pNR-LIB (Clontech); Site:1:		
	Still (ggcgcctcgcc); Site:2: Still (ggccatcagcc);		
	Double-stranded cDNA was prepared from cell line RNA. 5'		
	and 3' adaptors were used in cloning as follows: 5'		
	adaptor sequence: 5'-CACGCCATTAGGCC-3' and 3' adaptor		
	sequence: 5'-ATTCTAGAGCCGAGCGGCCGACACAG-dT(30)BN-3'		
	(where B = A, C, or G and N = A, C, G, or T). Average		



QY 145 ProbenValCysPheHisValSerPheAlaAspAspAlaGluGlyLysAlaIleIle 164  
 DB 407 GATTAATTAATATTTT---GTAATCTCTGTTCTCCATGACAGAAAGAAAAGATTAT 463  
 QY 165 AspGlnIleValAlaLeuAsnSerGly-----SerValIleuValAspGlyLeu 180  
 DB 464 GATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 523  
 QY 181 GlnIleuGlnIleuProAlaValCys-----GlnGluPhe 192  
 DB 524 AAATAGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 583  
 QY 193 ValAsnSerValPheCysGlnGluGlnIleuValThrGluValValVal 210  
 DB 584 CTCCTCTCTGTTCTTGAAGAGAACTATTAATTAATTAATTAATTAATTAATTAAT 637  
 RESULT 11  
 CNS03H8Z/c 985 bp DNA linear GSS 17-MAY-2000  
 LOCUS Tetradon nigroviridis genome survey sequence T7 end of clone  
 DEFINITION 026E24 of library G from Tetradon nigroviridis, genomic survey  
 sequence.  
 ACCESSION AL243980.1 GI:7964992  
 VERSION AL243980  
 KEYWORDS GSS; genome survey sequence.  
 SOURCE Tetradon nigroviridis.  
 ORGANISM Tetradon nigroviridis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Acanthomorpha; Acanthopterygii; Teleostei; Euteleostei; Neoteleostei;  
 Tetraodontidae; Tetraodon.  
 1 (bases 1 to 985)  
 Roest-Crollius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,  
 Bernot, A., Fizames, C., Wincker, P., Broctier, P., Quecier, F.,  
 Saurin, W. and Weissenbach, J.  
 Human gene number estimate provided by genome wide analysis using  
 Tetradon nigroviridis DNA sequence  
 Unpublished  
 2 (bases 1 to 985)  
 Roest-Crollius, H., Jaillon, O., Dasilva, C., Fizames, C., Fisher, C.,  
 Bouneau, L., Billault, A., Quecier, F., Saurin, W., Bernot, A. and  
 Weissenbach, J.  
 Characterization and repeat analysis of the compact genome of the  
 freshwater pufferfish Tetradon nigroviridis  
 Unpublished  
 3 (bases 1 to 985)  
 Genoscope.  
 Direct Submission  
 Submitted (12-APR-2000)  
 This sequence is a single read and was generated as part of a large  
 scale clone-end-sequencing project of the Tetradon nigroviridis  
 genome. For more information, please take a look at  
 http://www.genoscope.cns.fr/Tetraodon.  
 Location/Qualifiers  
 source  
 1. 985  
 /organism="Tetradon nigroviridis"  
 /db\_xref="taxon:99883"  
 /clone="026E24"  
 /clone\_1lb="G"  
 /note="Genoscope sequence ID : COBG026BC12LPI-end : T7"  
 BASE COUNT 246 a 269 c 238 g 230 t 2 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 3.3 Length: 985  
 Score: 86.00 Matches: 45  
 Percent Similarity: 36.62% Conservative: 33  
 Best Local Similarity: 21.13% Mismatches: 59  
 Query Match: 7.69% Indels: 76  
 DB: 17 Gaps: 12  
 US-10-034-500-2 (1-218) x CNS03H8Z (1-985)

QY 19 AspTyrSerGly-----SerMetMetLysHisValAlaValArg---Glu 33  
 DB 732 GACTATCTGTTTTCAGCGCTGCGGTTTATTAACATGACACACACCTACAGACAGAC 673  
 QY 34 ProValIleGluLeuAlaLysGluAlaIleLeuValIleAsnAlaIleMetProLysMet 53  
 DB 672 CTGAAGATCTAGCGGCTGACAGAGGAGGTGAGAGACGGCTGCA----- 625  
 QY 54 SerTyrGlnGlyLeuTyrThrPheAlaProTyrSerValIleIleProGlnIlySer 73  
 DB 624 -----GCCCTTGCTAAGGTTTGTCGCTCTTTCCTCATGCTAGC 586  
 QY 74 Tyr--AsnSerCysValAlaGluCysAlaValAsnThrIleLysSerAspLeuGluIle 92  
 DB 585 TGGGAGAGTACCGTCACATGACGCTGTTGTTGTTCTCCCTGACAGGCTTCGCTCG 526  
 QY 93 PheGlyArgLeuThrProVal-----ValAspGlyIle----- 99  
 DB 525 GAGGCGAGCTGACGCCATCTGTCGACATGTGTAAGAGCGCCAACTAGACGGCTG 466  
 QY 100 -----GlyAspGlyIle-----LysMetHis 106  
 DB 465 CTGACCAACACACAGCGACTCTCTGACAGCTGCCAACACCGCTCAAGCGGCTCAC 406  
 QY 107 GluThrValIleAsnGlnMetProProGlnAlaValIleLeuThrAspGlyHis 126  
 DB 405 GAG-----ATCTGCAGACCGAC----- 388  
 QY 127 AsnAsnLeuGlyMetAsnProValGlnGluValLysSerIleTyrGlnThrAsnProAsn 146  
 DB 387 -----AAGGCTTCACGACGATGAACTACACACAGGATTAATTAATTAATTAAT 334  
 QY 147 -----ValCysPheHisValSerPheAlaAspAsp 157  
 DB 333 CTAAAGCTGTGTAACATTAATAATTAATGCTTTATGATGATGACGTGCTT----- 280  
 QY 158 AlaGluGlyLysAlaIleIleAspGlnIleValAlaLeuAsnSerGlySerValLeuVal 177  
 DB 279 -----TGATTTCTGTGACAGCTGCGACCAACACAAACGCGCTCG---TTGGTG 235  
 QY 178 AspGlyLeuGlnLeuLeuGlnAsnPro---AlaValCys 189  
 DB 234 AATTCATGAGATGATTCAGAACCCGCTGCGACCTGT 196  
 RESULT 12  
 LOCUS BG491140 1411 bp mRNA linear EST 27-MAR-2001  
 DEFINITION 602519984F1 NIH\_MGC\_18 Homo sapiens cDNA clone IMAGE:463405 5',  
 mRNA sequence.  
 ACCESSION BG491140  
 VERSION BG491140.1 GI:13452652  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE NIH-MGC http://mgi.nci.nih.gov/  
 1 (bases 1 to 1411)  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 CONTACT: Robert Strauberg, Ph.D.  
 Email: c9abs-remail.nih.gov  
 Tissue Procurement: DCD/DTP/Gazdar  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL ac:  
 http://image.llnl.gov  
 plate: LCM400 row: h column: 22  
 High quality sequence start: 14  
 High quality sequence stop: 49.  
 Location/Qualifiers  
 FEATURES

source

1. 1411  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4638405"  
/clone\_1db="NIH\_MGC\_18"  
/tissue\_type="large cell carcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

BASE COUNT 378 a 316 c 376 g 341 t  
ORIGIN

Alignment Scores:  
Pred. No.: 6.95 Length: 1411  
Score: 85.50 Matches: 34  
Percent Similarity: 44.25% Conservative: 16  
Best Local Similarity: 30.09% Mismatches: 40  
Query Match: 7.64% Indels: 23  
DB: 12 Gaps: 7

US-10-034-500-2 (1-218) x BG491140 (1-1411)

Qy 4 ThrAlaSerCythrIlyAArgValGluSerTyraSerTyrluValAspTyrSerGlySer 23  
Db 748 ACAGATTCGTGATCCAAAGATTGAACTTAT--TACGTGCGCCATATAGGAAC 804

Qy 24 MetMetMetIlyHisValAlaValArgGlu-----ProlyIleGluLeuAla 39  
Db 805 TGTATGCTCGGCGATGCCCAAGTCACCAATGCCCCCTGAT--TTGGCG 861

Qy 40 LysGluAlaIleLeuIlyIleAsnAlaAlaMetProlyMetSerTyrlngIlyGluLeu 59  
Db 862 GGAAAGGCGATC-----GCATTGCGGAAAGTTATATTAACAACACACAC 906

Qy 60 TyrThrPheAlaProTyrSerValIleIleProGlnGlySerTyrPheSerCysValAla 79  
Db 907 CATAGGTTGGCCCACTCTCGCT-----TGGAAATTTTGTGTCG 948

Qy 80 GlyCysAlaValAsnThrIleIlySerAspLeuGluIlePheGlyArgLeuThrProVal 99  
Db 949 CAGAGCGCTCCATGCA--AAAGACAATCTATAGCACTTAGGTTTG----- 993

Qy 100 GlyAspGlyIleIlyMetHisGluThrValIleAsnGln 112  
Db 994 ---GACGTGTTAGAGGGAGACACTTTCGTGAGCAA 1029

RESULT 13  
LOCUS AZ203552 917 bp DNA linear GSS 31-AUG-2000  
DEFINITION SP 0061 B2 A08 SP6E Strongylocentrotus purpuratus, purple sea urchin, sperm genomic BAC library Strongylocentrotus purpuratus genomic clone Plate=61 Col=16 Row=B, DNA sequence.  
ACCESSION AZ203552  
VERSION AZ203552.1 GI:8398470  
SOURCE GSS.  
ORGANISM Strongylocentrotus purpuratus.  
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa; Echinoidea; Euechinoidea; Echinacea; Echinoda;  
Strongylocentrotidae; Strongylocentrotus.  
REFERENCE  
AUTHORS Cameron,R.A., Mahlitz,G., Rast,J.P., Martinez,P., Biondi,T.R., Swartzell,S., Wallace,J.C., Roustka,A.J., Livingston,B.T., Wray,G.A., Etenelson,C.A., Lentsch,H., Britten,R.J., Davidson,E.H. and Hood,L.  
TITLE A sea urchin genome project: Sequence scan, virtual map, and additional resources

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (17), 9514-9518 (2000)  
MEDLINE 20402566  
COMMENT Contact: Cameron, RA, Davidson, EH, Hood, L  
Division of Biology 156-29  
California Institute of Technology  
Pasadena California 91125, USA  
Tel: (626) 395-8421  
Fax: (626) 793-3047  
Email: acameron@caltech.edu  
Plate: 81 row: B column: 16  
Seq primer: Sp6  
Class: BAC ends  
High quality sequence stop: 917.  
Location/Qualifiers  
1. 917  
/organism="Strongylocentrotus purpuratus"  
/db\_xref="taxon:7668"  
/clone="Plate=81 Col=16 Row=B"  
/clone\_1db="Strongylocentrotus purpuratus, purple sea urchin, sperm genomic BAC library"  
/note="Organ: sperm; Vector: BACs 3.6; BAC clones in E-Coli DH10B"

BASE COUNT 218 a 211 c 205 g 278 t 5 others  
ORIGIN

Alignment Scores:  
Pred. No.: 3.95 Length: 917  
Score: 85.00 Matches: 32  
Percent Similarity: 40.65% Conservative: 18  
Best Local Similarity: 26.02% Mismatches: 41  
Query Match: 7.60% Indels: 32  
DB: 17 Gaps: 7

US-10-034-500-2 (1-218) x AZ203552 (1-917)

Qy 64 ProTyrSerValIleIleProGlnGlySerTyrPheSerCysValAlaGluCysAlaVal 83  
Db 75 CCGNTTAAGGTGATATGTCATTAGCATGTGTCGCGCAGATGTGTGTTGCA 134

Qy 84 AsnThrIleIlySerAspLeuGluIlePheGlyArg-----LeuthProValGlyasp 101  
Db 135 AATTCCTTAGAC-----ATCGGTGTACGACATACCTCT----- 173

Qy 102 GlyIleIlyMetHisGluThrValIleAsnGlnMetProProGlnAlaAlaValIleLeu 121  
Db 174 -----CATGAAGCGGTGCTTCTTGCGCCCAACATGGGGTTC----- 212

Qy 122 LeuthrAspGlyHisAsnAsnLeuGluIlyMetAsnProValGluGluValIlySerIleTyr 141  
Db 213 -----TTGGCAACCTTGT-----GAGCTGTGATGCTGTAC 245

Qy 142 GlnThrAsnProAsnValCysPheHisValIleSerPheAlaAspAspAlaGluGlyLys 161  
Db 246 TATCTCGCTCCAAATATGTGTGTGTCGCGCATGTATGATGATGATGAG----- 299

Qy 162 AlaIleIleAspGlnIleValAlaLeuAsnSerGlySerVal-----LeuValasp 178  
Db 300 -----GATGGCAGCGCAATTCAGCTCTCTCCCGCTGAGCCAGCAAGATTCTTGA 353

Qy 179 GlyLeuGln 181  
Db 354 GGTGTTCAA 362

RESULT 14  
LOCUS BM246457 713 bp mRNA linear EST 01-FEB-2002  
DEFINITION K0738D11-3 NTA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1-) cDNA library (long) Mus musculus cDNA clone K0738D11 3', mRNA sequence.  
ACCESSION BM246457  
VERSION BM246457.1 GI:17881727  
KEYWORDS EST.  
SOURCE house mouse.

ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	1 (bases 1 to 713)
TITLE	Piao, Y., Karyul, G.D., Dudekula, D.B., Qian, Y., Luo, A., Carter, M.G., Alba, R., Taub, D., Longo, D.L., Keller, J. and Ko, M.S.H. Systematic Analyses of NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1-) cDNA Library (Long)
JOURNAL	Unpublished (2001)
COMMENT	Contact: Dawood B. Dudekula Laboratory of Genetics National Institute on Aging/National Institutes of Health 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA Email: cdna@gsun.grc.nia.nih.gov Plate: K0738 row: D column: 11 Seq primer: -21M13 Forward High quality sequence stop: 713 POLY=A=yes.
FEATURES	Location/Qualifiers
SOURCE	1..713 /organism="Mus musculus" /strain="C57BL/6NCR" /db_xref="niaST:K0738D11-3" /db_xref="taxon:10090" /clone="K0738D11" /clone_1ib="NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1-) cDNA library (Long)" /tissue_type="Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1-) "/ /dev_stage="Age approx.10 weeks old" /lab_host="MDH10B" /note="Vector: pSPORT1 (Invitrogen); Site_1_Sali; Site_2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH ( <a href="http://ligsun.grc.nia.nih.gov/cdna">http://ligsun.grc.nia.nih.gov/cdna</a> ). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were obtained from Drs. Dennis Taub, Dan Longo (National Institute on Aging, USA), Jonathan Keller (National Cancer Institute, USA). Double-stranded cDNAs were synthesized with an Oligo(dT) primer (Invitrogen): 5'-GCAGACGATTCAGATCGGACGGCCCTTTTGTTTT-3' from 0.9 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker Lx-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Tag polymerase (Takara) with a primer Sal4-5. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with Sali and NotI enzymes and cloned into Sali/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.1 kb. The library was constructed by Yulan Piao (NIA)."
BASE COUNT	214 a 139 c 115 g 245 t
ORIGIN	
Alignment Scores:	
Prod. No.:	3.03 Length: 713
Score:	84.50 Matches: 30
Percent Similarity:	43.24% Conservative: 34
Best Local Similarity:	20.27% Mismatches: 69
Query Match:	7.55% Indels: 15
DB:	13 Gaps: 4
US-10-034-500-2 (1-218) x BM246457 (1-713)	
Gy	51 ProlyMetSerTyf-----glnGlyLeuTyThrPheAlaProtyIser 66
Dd	607 CCTAAATCAACTATTCTTGTCGAAGAAGAGAAATATTGTGTAAGCATGCTTTACT 548
Gy	67 ValIetIePrGInGlySerTrpAenSerCySValAlaGluCyaAlaValanThrIle 86

	Bd	547	GTAAGTAAAGAAAGAAGAAATGGACCTACATTTGCTATAGATAGAGATGATACAGG	488
	Oy	87	LysSerAspLeuGlulIlePheGlyArgLeuThrProVal-----GlyAspGlyIleLys	104
	Db	487	ATGATGAAGAGTGAGGTCTTAGCGACGACTGCACAATATTGCTTGAATCTGGAGATAA	428
	Oy	105	MethIseGluThrValIleAsnGlnMetCProProGlnAlaIleValIleLeuLeuThrAsp	124
	Db	427	CTTAGACTCATATATGCTTTAAACGTATCCCAACAGAAAAAACACAGCTAGATCTACA	368
	Oy	125	GlyHisAsnAsnLeuGlyMetAsnProValIGlulGluValIlySerIleTyrl-n-----	142
	Db	367	ACAACAGTACATACATGACGAGSTTATCAAAGCTAAGAAATTAAGAAAACACCACCTCACTCTG	308
	Oy	143	-----ThraProAsnValCysePheHISValIleSerPheAlaAspAspAla	158
	Db	307	CTTCCTTATGAACATGTCATTAATCTTTGGCGGACTTGACATCATGATGTTTAT	248
	Oy	159	GIuGLylYsaIalIeIleAsp-----GlnIleValAlaIleuAsnScrGlySerVal	175
	Db	247	AAAGATAGATCTCTGCCAATAAGAAAAATGTCATTCACTGATTAAGCAAGCTCT	188
	Oy	176	LeuValAspGlyLeuGlnLeuLeu	183
	Db	187	TTGGTTTAGGAATGGCACCTTATT	164
RESULT 15				
LOCUS	BOS18708/c		801 bp	mRNA linear EST 10-JUN-2002
DEFINITION	BSF626123		Generation of a set of potato cDNA clones for microarray analyses mixed potato tissues Solanum tuberosum cDNA clone STMJL66	
ACCESSION	BOS18708			
VERSION	BOS18708			
KEYWORDS	BOS18708.1 GI:2137577			
SOURCE	EST.			
ORGANISM	potato.			
REFERENCE	Solanum tuberosum			
AUTHORS	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum.			
TITLE	Buell,C.R., Harr,A., Baker,B., Tanksley,S., Fry,W., Smart,C., Rastebro,S., Griffiths,H., van der Hoeven,K., Tsai,T. and Karamycheva,S.A.			
JOURNAL	Generation of a set of potato cDNA clones for microarray analyses Unpublished (2002)			
COMMENT	Other ESTs: EST626122 Contact: Robin Buell The Institute for Genomic Research 9712 Medical Center Dr, Rockville, MD 20850, USA Email: potatoc@tigr.org This clone is available through the Research Genetics, contact the Research Genetics for further information 1-800-711-6195 or cdna@esgen.com Seq primer: T7. Location/Qualifiers 1..801 /organism="Solanum tuberosum" /cultivar="Kennebec Or Binjle" /db_xref="taxon:4113" /clone="STMJL66" /clone_id="Generation of a set of potato cDNA clones for microarray analyses mixed potato tissues" /issue_type="mixed tissues" /lab_host="SOLR" /note="Vector: pBluescript SK(-), Site_1: EcoRI, Site_2: XhoI, supplier: Combination of untreated and Phytophthora infestans-treated libraries of scions, leaves, leaflets, axillary buds of stem explants, petioles, germinating eyes 'tubers', or roots."			
BASE COUNT	228 a	161 c	212 g	200 t
ORIGIN				

## Alignment Scores:

Pred. No.: 3.67 Length: 801  
 Score: 84.50 Matches: 33  
 Percent Similarity: 38.03% Conservative: 21  
 Best Local Similarity: 23.24% Mismatches: 65  
 Query Match: 7.55% Indels: 23  
 DB: 14 Gaps: 5

US-10-034-500-2 (1-218) x BQ518708 (1-801)

Qy 8 ThrlySArgValGluSerTyrAsn---TyrLeuValAspTyrSerGlySerMetMet 26  
 Db 726 ACCAATAGCTGCAGCTTATATCAACATCTTGCTGCTTCCCTGCT----- 679  
 Qy 27 LysHsValAlaValArgLupProLysIleGluLeuAlaLysGluAlaIleLeuLysIle 46  
 Db 678 -----CAAGACCCCGCCCTACCTATGATGCCACCTCTGCTGCTGCT 634  
 Qy 47 AsnAlaAlaMetProLysMetSerTyrGlnGlyGlyLeuTyrThrPheAlaProTyrSer 66  
 Db 633 CCAGAGCGCATGCCCCCACAATTAATGGACAGTCCAGACCTCCCATTTGCCAATTCCA 574  
 Qy 67 ValIleIleProGlnGlySerTyrPasnSerCysValAlaGluCysAlaValAsnThrIle 86  
 Db 573 GTTCTTGCTGCTCAGGT-----TATCCAGCGGTT 544  
 Qy 87 LysSerAspLeuGluIlePheGlyArgLeuThrProValGlyAspGlyIleLysMetHis 106  
 Db 543 CCTCCAATGTCACCAATGTCAGGCGCAGATGCCACCTGTGCA---TCCTGCGGATGCAA 487  
 Qy 107 GluThrValIleAsnGlnMetProProGlnAlaAlaValIleLeuLeuThrAspGly--- 125  
 Db 486 CCCAATGCTCTTCAGCTCTCTCTGCACGTAAACCCCTTGTAGAGTTCCAGGTGAACC 427  
 Qy 126 ---HisAsnAsnLeuGluIleMetAsnProValGluGluValIleSerIleTyrGlnThrAsn 144  
 Db 426 CCTCATCTATATCTAGGACAGCGCCCTTTCTGCTACATCAATCATATGCCAGCAAT 367  
 Qy 145 ProAsn 146  
 Db 366 CCCAAT 361

RESULT 16  
 LOCUS BQ792380 835 bp mRNA linear EST 30-JUL-2002  
 DEFINITION EST 8835 Veralison Grape berries Lambda Triplex2 Library Vitis  
 vinifera cDNA clone CT006B11 3', mRNA sequence.  
 ACCESSION BQ792380  
 VERSION BQ792380.1 GI:22007346  
 KEYWORDS EST.  
 SOURCE Vitis vinifera.  
 ORGANISM Vitis vinifera.  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Vitaceae; Vitis.  
 1 (bases 1 to 835)  
 Abbal,P., Agase,A., Ageorges,A., Atanaseva,R., Barrieu,F.,  
 Couture,C., Dedalechamp,F., Delrot,S., Gilsant,D., Grimplet,J.,  
 Hamdi,S., Romeu,C. and Texier,N.  
 Generation of Expressed Sequence Tag from Grape Berry (skin, pulp  
 or seeds) at Various Developmental Stages  
 Unpublished (2002)  
 Contact: Hamdi S.  
 UMR 619 - Equipe Biologie de la Vigne  
 Universite de Bordeaux I, Institut National de la Recherche  
 Agronomique  
 71, Avenue Edouard Bourlaux, BP 81, 33883 Villenave D'Ornon Cedex,  
 France  
 Tel: 00-33-(0)5-57-12-25-50  
 Fax: 00-33-(0)5-57-12-25-48  
 Email: s.hamdi@bordeaux.inra.fr  
 Seq primer: T7.

FEATURES  
source

Location/Qualifiers  
 1..835  
 /organism="Vitis vinifera"  
 /cultivar="Cabernet Sauvignon"  
 /db\_xref="taxon:29760"  
 /clone="CT006B11"  
 /clone\_id="Veralison Grape berries Lambda Triplex2  
 Library"  
 /dev\_stage="veraison stage"  
 /note="Organ: Fruit without seeds; Vector: Lambda Triplex2  
 ; Site\_1: SfiIA, Site\_2: SfiIB; Oriented library"

BASE COUNT 245 a 202 c 143 g 245 t  
 ORIGIN

## Alignment Scores:

Pred. No.: 3.93 Length: 835  
 Score: 84.50 Matches: 43  
 Percent Similarity: 38.50% Conservative: 29  
 Best Local Similarity: 22.99% Mismatches: 60  
 Query Match: 7.55% Indels: 55  
 DB: 14 Gaps: 9

US-10-034-500-2 (1-218) x BQ792380 (1-835)

Qy 12 GluSerTyrAsnTyr-----LeuValAspTyrSerGlySerMet 24  
 Db 704 GAGCATTTCACTACTCTTGCAGAGTAGAAGTTCCTCTGTGTGCTATGTCAGAGAGATGCT 645  
 Qy 25 MetMetLys-----HisValAlaValArgLupPro 34  
 Db 644 TTATTGACATATTGATACACCCTCTTGTGAGTCATCTCAGCTGATATGCTAGAGCA 585  
 Qy 35 LysIleGluLeuAlaLysGlu-----AlaIleLeuLysIleAsnAlaAla 49  
 Db 584 AAGCGAGATATATGCTTCACTGATGAACACCTCTGGCCATTCGTAATACAGAAATG 525  
 Qy 50 MetProLysMetSerTyrGlnGlyLeuTyrThrPheAlaProTyrSerValIleIle 69  
 Db 524 CTTTTTTATGACACTGCTGAGGAGGAGTACTGCTGCGGCATAC----- 477  
 Qy 70 ProGlnGlySerTyrPasnSerCysValAlaGluCysAlaValAsnThrIleLysSerAsp 89  
 Db 476 -----TGTCCAAAGCAGTGTGAGGTCTGTCT---AATGTTGCCAATCACAG 432  
 Qy 90 -----LeuGluIlePheGlyArgLeuThrProValGlyAspGlyIleLysMetHis 106  
 Db 431 GCAGACATCTTGAATCTCCCCCAGAGACTCAAAAGAGAGGTAAAGATGCTA 372  
 Qy 107 GluThrValIleAsnGlnMetProProGlnAlaAlaValIleLeuLeuThrAspGlyHis 126  
 Db 371 CTGTATCATTTGAAT-----GTGCCAGCACAGAGGTATATGCTATGCTATGAGAGAA 318  
 Qy 127 AsnAsnLeuGlyMet-----AsnProValGlu 135  
 Db 317 AATGACATTTGAATGTTGAATTAAGCTTTTAGGCATTCTCTCAATATGCTTCGGAG 258  
 Qy 136 GluValLysSerIleTyrGlnThrAsnProAsnValCysPheHisValLysSerPheAla 155  
 Db 257 AACACAAAGCTGTG-----GCTAATGTATTTGTGCTCCAGC 222  
 Qy 156 AspAspAlaGluGlyLysAla 162  
 Db 221 AATGACGAGAGATGTGTACT 201

RESULT 17  
 LOCUS B1611208 588 bp mRNA linear EST 07-SEP-2001  
 DEFINITION RH17375.Sprime RH Drosophila melanogaster normalized Head pRLC-1.  
 Drosophila melanogaster cDNA clone RH17375 5 similar to CG5903:  
 FBan005903 GO: [ ] located on: 3R 89B5-89B9; 08/17/2001, mRNA  
 sequence.  
 ACCESSION B1611208  
 VERSION B1611208.1 GI:15506733



KEYWORDS EST.  
SOURCE fruit fly.  
ORGANISM Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
REFERENCE  
1 (bases 1 to 588)  
Stapleton, M., Brokstein, P., Hong, J., Tyler, D., Berman, B., Carlson  
, D., Chame, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George  
, R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Miera, S.,  
Mingall, C.J., Nuno, J., Paclob, J., Paragae, V., Park, S.,  
Phouanavong, S., Wan, K., Yu, C., Lewis, S.E., Celniker, S. and Rubin  
, G.M.  
BDGP/HMT RH Drosophila EST Project  
TITLE  
JOURNAL Unpublished (2001)  
COMMENT Contact: Stapleton, M.  
BDGP  
Lawrence Berkeley National Lab  
One Cyclotron Rd, Berkeley, CA 94720, USA  
Fax: 510 486 6798  
Email: <http://www.fruitfly.org/EST/estfruitfly.berkeley.edu>  
hit genomic AB003712: arm:3R [11875180,12101379]  
estimated-ctto:89B-89B21: 08/17/2001  
Plate: RH.173 row: G column: 3  
High quality sequence stop: 446.  
Location/Qualifiers  
1. 588  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"  
/clone="RH17375"  
/clone\_1b="RH Drosophila melanogaster normalized Head  
p1c-1"  
/sex="male and female"  
/dev\_stage="adult"  
/lab\_host="DHS-alpha Tona"  
/note="Organ: head, Vector: p1c1, Site: 1: XhoI; Site: 2:  
BamHI; Library was kindly generated by Piero Carninci at  
the RIKEN. The library was normalized and excised using  
Cre recombinase. Plasmid cDNA library."  
BASE COUNT 142 a 179 c 160 g 107 t  
ORIGIN

Alignment Scores:  
Pred. No.: 2.56 Length: 588  
Score: 84.00 Matches: 39  
Percent Similarity: 44.88% Conservative: 18  
Best Local Similarity: 30.71% Mismatches: 40  
Query Match: 7.51% Indels: 30  
Gaps: 8

US-10-034-500-2 (1-218) x B1611208 (1-588)

Qy 76 SerCySvAlaIaGluCyAlaValaAnthrIlelySserApLeuGluIlePheGlyArg 95  
Db 246 TCCAGTGGCGGACTGAGCTGTGTGCCGCCGCCGAGCTCCATCTACGGCAGC 305  
Qy 96 LeuthrProValaGlyAspGlyIlelySvethIsgIuThrValIleAsnGlnMeProPro 115  
Db 306 TTGGCGCAAGACAAACCAAGCCAGCGGCT-----CCACCA 344  
Qy 116 GlnAlaAlaValaIleLeuLeuThraSpGlyHIsaMenLeu-----GlyMetAsnPro 133  
Db 345 CAAAGCTGTGTGTG-----CACAAAGACCTGTGAGCGCGGAGTCCGCTAT 389  
Qy 134 Val---GluGluValaValSerIleTyrlGlnThr-----AsnProAsnValCyaspHe 149  
Db 390 GTGGCGGAGGAGTGTGATGTGATTAAGCAAGTGCAGTCAAGCCGCAATGTCGCGC 449  
Qy 150 HisValValSerPheAlaAspAlaAspAlaGluGlyLysAlaIleIleAsnGlnIleValAla 169  
Db 450 CATTAACGTGGAACCGCT---AAGCACACACCCAGTCCACATGACATG-----497  
Qy 170 LeuAen-----SerGlySerValLeuValaIleValaIleGln 181

Db 498 CTGAACGAGCCCAAGAACTCCCTGACCGCAGCGAGCCATTGTGGCGGCTGGCT 557  
Qy 182 LeuGlnAenProAlaVal 188  
Db 558 GGTATATCTTGTGCGCGCTG 578

RESULT 18  
LOCUS BM911554  
DEFINITION  
AGENCOURT\_6615185 NIH\_MGC\_41 Homo sapiens cDNA clone IMAGE:5466355  
5', mRNA sequence.  
ACCESSION  
BM911554  
VERSION  
BM911554.1 GI:19361933  
KEYWORDS  
EST.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 1144)  
NIH-MGC <http://imgc.ncl.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: DCTD/DTF  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.lnl.gov>  
Plate: LICM1966 row: j column: 20  
High quality sequence stop: 597.  
Location/Qualifiers  
1. 1144  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5466355"  
/clone\_1b="NIH\_MGC\_41"  
/issue\_type="amelanotic melanoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: skin; Vector: pOT87; Site: 1: XhoI; Site: 2:  
EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGACACAG(C). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH MGC library."  
BASE COUNT 240 a 352 c 366 g 185 t 1 others  
ORIGIN

Alignment Scores:  
Pred. No.: 7.7 Length: 1144  
Score: 84.00 Matches: 30  
Percent Similarity: 48.42% Conservative: 16  
Best Local Similarity: 31.58% Mismatches: 34  
Query Match: 7.51% Indels: 15  
Gaps: 5

US-10-034-500-2 (1-218) x BM911554 (1-1144)

Qy 12 GluSerTyrlAenTyrlLeuValaAspTyrlSerGlySvethMetClyGHisValAlaVal 31  
Db 565 CAGTCGTACCGGGGTGTGTGCTACGTACGACGACGAGTGCAGGACAGTACGCTG 624  
Qy 32 ArgGluProLysIleGluLeuAlaValGluAlaIleLeuLysIleAenAlaIleMePro 51  
Db 625 CGGAGCCCGACATCCGGAACGTGACAG-----CTCAAGGGAACCATCAG 672  
Qy 52 LysMetSerTyrlGlnGlyLeuTyrlThr-Phe-----AlaProTyrlSerValIi 68



Qy	Db	673	AGCCTGAGATGGAGGCGCCCTC--ACCCCTCGGGGGGGAGGCCCCCGCAACATC--	727
Qy	68	e1leProGInglYseRTIPaNsErCySValaJlucYsAlaValaAsnThrIleYsSe	88	
Db	728	-----TCGGAGAGCAGCGTGTGTGGCGCCCTCCCCCAACAACCGTTGCGCC	774	
Qy	88	rAspLeuGlulIlePheGly---ArgLeuThrProValGlyAap	101	
Db	775	CGGGTACTATTGTGACGGGCGCCTCCAGCACCCCTTAAGGGAGAT	817	
RESULT 19				
LOCUS	BE372805	625 bp	mRNA	linear
DEFINITION	60122432121 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3582797 5'			
ACCESSION	BE372805			
VERSION	BE372805.1			
KEYWORDS	EST.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
TITLE	1 (bases 1 to 625)			
JOURNAL	NIH-MGC http://mgc.ncl.nih.gov/.			
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)			
	Unpublished (1999)			
	Contact: Robert Strausberg, Ph.D.			
	Email: cga@bbs-r@mail.nih.gov			
	Tissue Procurement: Gilbert Smith, Ph.D.			
	cDNA Library Preparation: Life Technologies, Inc.			
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
	DNA Sequencing by: Incyte Genomics, Inc.			
	Clone distribution: MGC clone distribution information can be			
	found through the I.M.A.G.E. Consortium/LLNL at:			
	http://image.llnl.gov			
	Plate: LHAM8738 row: 1 column: 06			
	High quality sequence stop: 587.			
FEATURES	Location/Qualifiers			
source	1..625			
	/organism="Mus musculus"			
	/strain="FVB/N"			
	/db_xref="taxon:10090"			
	/clone="IMAGE:3582797"			
	/clone_lib="NCI_CGAP_Mam1"			
	/tissue_type="tumor, biopsy sample"			
	/dev_stage="10 months, virgin"			
	/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;			
	Site 2: NotI; Cloned unidirectionally. Primer: Oligo dr.			
	Library constructed by: Life Technologies. Investigator			
	providing samples: Gilbert Smith, NIH"			
BASE COUNT	171 a 132 c 155 g 167 t			
ORIGIN				
Alignment Scores:				
Pred. No.:	3.29	Length:	625	
Score:	83.50	Matches:	51	
Percent Similarity:	43.75%	Conservative:	26	
Best Local Similarity:	28.98%	Mismatches:	78	
Query Match:	7.46%	Indels:	21	
DB:	10	Gaps:	9	
US-10-034-500-2 (1-218) x BE372805 (1-625)				
Qy	2	GUUaUThraLaseCyThrLySArGValGluSeTyraSnTyLeuValaAspTySer	21	
Db	93	GAAGTGAAGCCCTCA-----MAGCGCTGGAGAAATTGTGTTCTTCACAGACCGTTCA	146	
Qy	22	GlySerMet-----MetMetLySHisValaAlaValaArgGluProLySileGluLeuAla	39	
Db	147	GGAAATATGACATCCCCCATGAGACAGAGAACAAATTCACGATACGCATCGAGCGTCC	206	
Qy	40	LySGluAlaIleLeuLySileAsnAlaAlaMeCProlYsMetSerTyR-----	55	

Db	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659
----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----







Db 422 CGGCTGCGCGGCTGCTCTCCGGCTGTTCAGAGCGACCGAGATCATGCGG----- 369  
 Qy 97 ThrProValGlyAspGlyIleIysMetHisGluThrValIleAsnGlnMetProProGln 116  
 Db 368 -----CTG 366  
 Qy 117 AlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsnProValGluGln 136  
 Db 365 GCGGATTCCTCTGTTGCACATGCGCGCAGC----- 333  
 Qy 137 ValIysSerIleTyrGlnThrAsnProAsnValCysPheHisValValSerPheAlaAsp 156  
 Db 332 -----ACCGCGCGGTTTGGGAC 315  
 Qy 157 AspAlaGluGlyLysAlaIleIleIleAspGlnIleValAlaLeuAsnSerGlySerValLeu 176  
 Db 314 GACGGTTCCGCTGATCTTCTGAGAGACAGATCGCTCTGTTGAAACGGCTTCATGATC 255  
 Qy 177 ValAspGlyLeuGlnLeuGlnAsnPro 186  
 Db 254 GCGGAAGCGCTGGCGCTGATCGACAGCGG 225  
 RESULT 25  
 CNS06H70 1007 bp DNA linear GSS 17-JUN-2001  
 LOCUS T3 end of clone AS0AA010D12 of library AS0AA from strain CLIB 533  
 DEFINITION of Saccharomyces bayanus, genomic survey sequence.  
 ACCESSION AL398650  
 VERSION AL398650.1 GI:12152318  
 KEYWORDS GSS.  
 SOURCE Saccharomyces bayanus.  
 ORGANISM Saccharomyces bayanus.  
 REFERENCE 1 (bases 1 to 1007)  
 AUTHORS Soucie,J.L., Aigle,M., Artiguenave,F., Blandin,G., Bolotin-Fukuhara,M., Bon,E., Broctier,P., Casaregola,S., deMontigny,J., Dujon,B., Durrens,P., Lepingle,A., Lorente,B., Malbertuy,A., Neuvéglise,C., Olier-Kalogeropoulos,O., Porter,S., Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M., Wincker,P. and Weissensbach,J.  
 TITLE Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies  
 JOURNAL FEBS Lett. 487 (1), 3-12 (2000)  
 MEDLINE 20584711  
 PUBMED 11152876  
 REFERENCE 2 (bases 1 to 1007)  
 AUTHORS Bon,E., Neuvéglise,C., Casaregola,S., Artiguenave,F., Wincker,P., Aigle,M. and Durrens,P.  
 TITLE Genomic exploration of the hemiascomycetous yeasts: 5. Saccharomyces bayanus var. uvarum  
 JOURNAL FEBS Lett. 487 (1), 37-41 (2000)  
 MEDLINE 20584715  
 PUBMED 11152880  
 REFERENCE 3 (bases 1 to 1007)  
 JOURNAL Genoscope.  
 TITLE Direct Submission  
 AUTHORS Submitted (07-SEP-2000) Genoscope - Centre National de Séquençage, 2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail : sequef@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
 COMMENT This GSS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces exiguus, Saccharomyces servazii, Zygosaccharomyces rouxii, Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila, Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.  
 FEATURES  
 source 1..1007  
 Location/Qualifiers  
 /organism="Saccharomyces bayanus"

/strain="CLIB 533"  
 /variety="uvarum"  
 /db\_xref="taxon:4931"  
 /clone="AS0AA010D12"  
 /clone\_1ib="AS0AA"  
 /note="End : T3"  
 complement(<104. .>388)  
 /note="similar to Saccharomyces cerevisiae ORF YDR203w ("  
 questionable ORF )"  
 /evidence=not\_experimental  
 <117. >1007  
 /note="similar to Saccharomyces cerevisiae ORF YDR202c ("  
 hypothetical protein )"  
 /evidence=not\_experimental  
 misc\_feature  
 BASE COUNT 357 a 185 c 205 g 258 t 2 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 9.78 Length: 1007  
 Score: 82.50 Matches: 45  
 Percent Similarity: 33.88% Conservative: 38  
 Best Local Similarity: 18.37% Mismatches: 75  
 Query Match: 7.37% Indels: 87  
 DB: 17 Gaps: 10  
 US-10-034-500-2 (1-218) x CNS06H70 (1-1007)  
 Qy 16 TyrLeuValAspTyrSerGlySerMetMetLeuHisValAlaValArgGluProLys 35  
 Db 384 TATATTATTGATTT-----CATATTGTGTAAATTTCACAA 422  
 Qy 36 IleGluLeuAlaGlyAlaIleLeuLysIleAsnAlaIleMetProLysMetSerTyr 55  
 Db 423 TTCAAAGGATTAACCAAGTATGTCGTATGATACCGGTTGAA----- 470  
 Qy 56 GlnGlyLysLeuTyrThrPheAlaProTyrSerValIleIleProGlnGlySerThrAsn 75  
 Db 471 -----TTCTTACTTGTCAGTCACTGTAATAATGACACAT----- 506  
 Qy 76 SerCysValAlaGluCysAlaValAsnThrIleLysSerAspLeuGluIle----- 92  
 Db 507 -----TGAAAAGTATTTGAAATATTGAACCA 536  
 Qy 93 -----PheGlyArgLeuThrProValGlyAspGlyIle 103  
 Db 537 TTCAACTGCTACAGATGTGAATGATGCGGAATAATTGACGTGCGCATGAGCTA 596  
 Qy 104 LysMetHisGluThrValIleAsnGlnMetProProGlnAlaIleValIleLeuLeuThr 123  
 Db 597 TTGAATCATTCCTCTGATCTCTACAAAATCCTCCAGAGAT-----CTGGTATTT 647  
 Qy 124 AspGlyHisAsnAsnLeuGlyMetAsnProValGluGluValIysSerIleTyrGlnThr 143  
 Db 648 CCAAGAGATTAACAATCAGCCATG-----AAGGAATGTTTCAAGT 689  
 Qy 144 AsnProAsnValCys-----PheHisValIleSerPhe----- 154  
 Db 690 TGTATTCAATCTGGGAATCTACGGCTCATATCTCGGACTGAAGTACGCTTGTAA 749  
 Qy 154 ----- 154  
 Db 750 AATGAGCTCTGATTGAATCAAGATCTATTAAAGTACAAAGAACTTGTGCGAA 809  
 Qy 155 AlaAspAspAlaGluGlyLysAlaIleIleAspGlnIle----- 167  
 Db 810 ATTGAATTAATAATCTGGCAAAATCATTTGTGACGATTAAGAAATCAAGTCAAAATGA 869  
 Qy 168 ValAlaLeuAsnSerGlySerValLeuValAspGlyLeuGlnLeuGlnAsnProAla 187  
 Db 870 AGACAAAGAACTTGTCCAAATCTTTTCAAGAAACGGTTCACAGTTACAGATTCCAG 929  
 Qy 188 ValCysGlnGluPheValAsnSerValPheCysGlnGluGlnIleLeuValThrGluGlu 207  
 Db 188 ValCysGlnGluPheValAsnSerValPheCysGlnGluGlnIleLeuValThrGluGlu 207

```

Db 930 CTACGATCATATATCTCTTCT---TTTCAAGTGAAGCATCACTCCGAGAGCC 986
Qy 208 ValValValLeuArg 212
   |||||
Db 987 CAGGCGCTATTAGA 1001

RESULT 26
BG463740
LOCUS EMI 51.D07.b1_A002 Embryo 1 (EMI) Sorghum bicolor cDNA, mRNA
DEFINITION BG463740
ACCESSION BG463740.1 GI:13391538
VERSION BG463740.1
KEYWORDS EST.
SOURCE sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 405)
AUTHORS Reid,S.P., Cordonnier-Pratt,M.-M., Gingle,A. and Pratt,L.H.
TITLE An EST database from Sorghum: developing embryos
JOURNAL Unpublished (2000)
COMMENT Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: emmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: JEN REV
High quality sequence stop: 388
POLYA=No.

FEATURES
    source
        1..405
            /organism="Sorghum bicolor"
            /db_xref="taxon:4558"
            /clone_id="Embryo 1 (EMI)"
            /note="Organ: Embryo germinated for 24 hr, Vector:
            pBluescript II from lambda Zap II, Site 1: XhoI, Site 2:
            EcoRI; The library was made from poly-A RNA in the cloning
            vector lambda Zap II. Clones to be sequenced were
            prepared by mass excision."
BASE COUNT 82 a 106 c 134 g 83 t
ORIGIN
Alignment Scores:
Pred. No.: 2.51 Length: 405
Score: 82.00 Matches: 23
Percent Similarity: 46.94% Conservative: 23
Best Local Similarity: 23.47% Mismatches: 44
Query Match: 7.33% Indels: 8
Gaps: 2
US-10-034-500-2 (1-218) x BG463740 (1-405)
Qy 114 ProProGlnAlaValAlIleuLeuThrAspGlyHisAsnAsnLeuGlyMetAsnPro 133
   |||||
Db 98 CTTGCGAAGCGCCAGGGTCCGCTGCACCGGCGCAAGAGGATCGGCTGAGAGTG 157
   |||||
Qy 134 ValGluGluValLeuSerIleTyrGlnThrAsnProAsnValCysPheHisValValSer 153
   |||||
Db 158 TGCCGCGACCTGCGCGC-----ACCGGGCTCACCGTCTCTTTAAACA 199
   |||||
Qy 154 PheAlaAspAspAlaGluGlyLysAlaIleLeuAspGlnIleValAlaLeuAsnSergly 173
   |||||
Db 200 GCCAAGGACGAGGCGGCGCTGCGGCGCTGGAAGAACTCGAGAGCTGGGCGCTTCC 259
   |||||
Qy 174 SerValLeuValAspGlyLeuGlnLeuLeuGlnAsnProAlaValCysGln-----Glu 191
   |||||

```

```

Db 260 GATGTCCTTTTCCACCAACTGAGACTTATCGATGCTCCGAGCATGCTCGAATTCGTAG 319
Qy 192 PheValAsnSerValPheCysGlnGluGlnIleLeuValThrGluGluValVal 209
   |||||
Db 320 TTCTCTAAGACCCGTTTCGGAGAGCTAGATATCTCGTGAATTAATGCGCGCATTT 373

RESULT 27
BE596201
LOCUS P11 50.G12.b1_A002 Pathogen induced 1 (P11) Sorghum bicolor cDNA,
DEFINITION BE596201
ACCESSION BE596201.1 GI:9851274
VERSION BE596201.1
KEYWORDS EST.
SOURCE sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 524)
AUTHORS Cordonnier-Pratt,M.-M., Gingle,A., Dean,R., Sudman,M. and Pratt
L.H.
TITLE An EST database from Sorghum: pathogen-induced plants
JOURNAL Unpublished (2000)
COMMENT Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: emmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: JEN REV
High quality sequence stop: 516
POLYA=No.

FEATURES
    source
        1..524
            /organism="Sorghum bicolor"
            /db_xref="taxon:4558"
            /clone_id="Pathogen induced 1 (P11)"
            /note="Organ: Anthracnose-infected leaves from
            two-week-old sorghum plants 48 hr after inoculation;
            Vector: pBluescript II from lambda Zap II; Site 1: XhoI;
            Site 2: EcoRI; Two-week-old sorghum plants (BTK 623
            cultivar) were infected with pathogen (isolate FRM421 of
            Colletotrichum graminicola, which is a sorghum isolate).
            RNA was prepared from infected leaves harvested from 45
            seedlings 48 hours after inoculation. Note: Young
            seedlings (2 weeks old) exhibit juvenile resistant
            reaction, which is an incompatible interaction. As they
            grow older (4 weeks or older), plants resume susceptibility
            to anthracnose disease. The library was made from poly-A
            RNA in the cloning vector lambda Zap II. Clones to be
            sequenced were prepared by mass excision. WARNING: While
            most or all ESTs are expected to derive from the host
            plant, no effort was made to eliminate ESTs deriving from
            the pathogen."
BASE COUNT 119 a 130 c 171 g 104 t
ORIGIN
Alignment Scores:
Pred. No.: 3.85 Length: 524
Score: 82.00 Matches: 23
Percent Similarity: 46.94% Conservative: 23
Best Local Similarity: 23.47% Mismatches: 44
Query Match: 7.33% Indels: 8
Gaps: 2
US-10-034-500-2 (1-218) x BE596201 (1-524)
Qy 114 ProProGlnAlaValAlIleuLeuThrAspGlyHisAsnAsnLeuGlyMetAsnPro 133

```

```

Db 79 CCTGGAGACGACGAGGCTCCGTCGTCACCGCGCAAGGATCCGGCTGAGAGTG 138
Qy 134 ValGluGluVallyseSerIleTyrgInThraNpRoAnValCyshPhetIValValser 153
Db 139 TCCCGGACAGTCCCGCC-----AACGGGTCACCGTCGTTTAAACA 180
Qy 154 PheaIAspAspAlaGluGlyLysAlaIleIleAspGlnIleValAlaLeuAsnSergly 173
Db 181 GCCAGGAGACGAGGCGCGCTCGCGCGGAGAACTCAGACCTCGGCGCTCTCC 240
Qy 174 SerValLeuValAspGlyLeuGlnLeuGlnAsnProAlaValCyseGln-----Glu 191
Db 241 GATGTCCTTTCCACCACTGACATTATCGATGCTCCAGACCTCGCAATTGCTGAG 300
Qy 192 PheValAsnSerValPheCysGlnGluGlnIleLeuValThrGluGluValVal 209
Db 301 TTCTCTCAAGACCGCTTCGGAGACTGATATCTGTGTAATATATGCGCGCAT 354

```

```

RESULT 28
BE599098 560 bp mRNA linear EST 18-AUG-2000
LOCUS P11_85.D01.b1.A002 Pathogen induced 1 (P11) Sorghum bicolor cDNA,
DEFINITION mRNA sequence.
ACCESSION BE599098
VERSION BE599098.1 GI:9854171
KEYWORDS EST.
SOURCE Sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 560)
Cordonnier-Pratt, M.-M., Gingle, A., Dean, R., Sudman, M. and Pratt
, L. H.
An EST database from Sorghum: pathogen-induced plants
Unpublished (2000)
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: JEN REV
High quality sequence stop: 523
POLYA=No.

```

```

FEATURES
source 1..560
Location/Qualifiers
/organism="Sorghum bicolor"
/db_xref="taxon:4558"
/clone_lib="Pathogen induced 1 (P11)"
/notes="Organ: Anthracnose-infected leaves from
two-week-old sorghum plants 48 hr after inoculation;
Vector: pBluescript II from lambda Zap II, Site 1: XhoI;
Site 2: EcoRI; Two-week-old sorghum plants (BRX 623
cultivar) were infected with pathogen (isolate PM421 of
Colletotrichum graminicola, which is a sorghum isolate).
RNA was prepared from infected leaves harvested from 45
seedlings 48 hours after inoculation. Note: young
seedlings (2 weeks old) exhibit juvenile resistant
reaction, which is an incompatible interaction. As they
grow older (4 weeks or older), plants resume susceptibility
to anthracnose disease. The library was made from poly-A
RNA in the cloning vector lambda Zap II. Clones to be
sequenced were prepared by mass excision. WARNING: While
most or all ESTs are expected to derive from the host
plant, no effort was made to eliminate ESTs deriving from
the pathogen."

```

```

BASE COUNT 127 a 142 c 177 g 114 t

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 4.3 Length: 560
Score: 82.00 Matches: 23
Percent Similarity: 46.94% Conservative: 23
Best Local Similarity: 23.47% Mismatches: 44
Query Match: 7.33% Indels: 8
DB: 10 Gaps: 2

```

US-10-034-500-2 (1-218) x BE599098 (1-560)

```

Qy 114 ProPProGlnAlaAlaValIleLeuLeuThraSpGlyHisAsnAsnLeuGlyMecAsnPro 133
Db 79 CCTGGAGACGACGAGGCGCGCTCGCGCGGAGAACTCAGACCTCGGCGCTCTCC 138
Qy 134 ValGluGluVallyseSerIleTyrgInThraNpRoAnValCyshPhetIValValser 153
Db 139 TCCCGGACAGTCCCGCC-----AACGGGTCACCGTCGTTTAAACA 180
Qy 154 PheaIAspAspAlaGluGlyLysAlaIleIleAspGlnIleValAlaLeuAsnSergly 173
Db 181 GCCAGGAGACGAGGCGCGCTCGCGCGGAGAACTCAGACCTCGGCGCTCTCC 240
Qy 174 SerValLeuValAspGlyLeuGlnLeuGlnAsnProAlaValCyseGln-----Glu 191
Db 241 GATGTCCTTTCCACCACTGACATTATCGATGCTCCAGACCTCGCAATTGCTGAG 300
Qy 192 PheValAsnSerValPheCysGlnGluGlnIleLeuValThrGluGluValVal 209
Db 301 TTCTCTCAAGACCGCTTCGGAGACTGATATCTGTGTAATATATGCGCGCAT 354

```

```

RESULT 29
BE585416 605 bp mRNA linear EST 16-AUG-2000
LOCUS 8-10B-20 Psoljae20 Phytophthora sojae cDNA, mRNA sequence.
DEFINITION BE585416
ACCESSION BE585416
VERSION BE585416.1 GI:9836365
KEYWORDS EST.
SOURCE Phytophthora sojae.
ORGANISM Phytophthora sojae
Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
Phytophthora.
1 (bases 1 to 605)

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Comparative analysis of expressed sequences in Phytophthora sojae
Quach, D., Hirber, P. T., Sobral, B. W. S. and GiJzen, M.
Plant Physiol. 123 (1), 243-254 (2000)
20267956
Contact: GiJzen M
Agriculture and Agri-Food Canada
1391 Sandford Street, London, Ontario, Canada N5V 4T3
Tel: 519 457 1470
Fax: 519 457 3997
Email: gijzenm@agr.ca.

```

```

FEATURES
source 1..605
Location/Qualifiers
/organism="Phytophthora sojae"
/strain="race 2, strain P6487"
/db_xref="taxon:67593"
/clone_lib="Psoljae20"
/dev_stage="zoospores"
/lab_host="E. coli strain XL0LR"
/notes="Vector: pBK-CMV; Site 1: EcoRI; Site 2: XhoI; This
cDNA library was constructed from polyA+ enriched mRNA
from zoospores grown in liquid medium. Zoospores were
released into water and collected by centrifugation at
2,000g; zoospore-bearing sporangia were induced to
develop on 5 to 7 d old mycelium colonies grown on V8
agar by repeated flooding with water. Complementary DNA
was synthesized from mRNA using an XhoI-poly(dT)
linker-primer. EcoRI adapters were ligated to the
blunt-ended cDNA fragments and the products were digested
with XhoI for directional cloning into lambda Zap Express

```



vector. This lambda library was amplified once using E. coli host strain XL1 Blue MRF<sup>+</sup>. Inserts were then subcloned by mass excision using Exsist helper phage for conversion into phagemid vector PBK-CMV in E. coli host strain XL0R. Sequenced using 3 primer: 5' ATT AAC CCT CAC TAA AGG CA 3'.

BASE COUNT 105 a 166 c 193 g 140 t 1 others  
ORIGIN

## Alignment Scores:

Pred. No.:	4.89	Length:	605
Score:	82.00	Matches:	48
Percent Similarity:	40.61%	Conservative:	32
Best Local Similarity:	24.37%	Mismatches:	73
Query Match:	7.33%	Indels:	44
DB:	10	Gaps:	8

US-10-034-500-2 (1-218) x BE585416 (1-605)

```

Qy 44 LeuYslleAaAlaAlaMeEProLyMeSeTyrgInGlyGly----- 58
Db 53 CTCTGATCACTAGTCTGCTTCTCGGTGCTGACTTGGAGTTTCGCTGGTGATCG 112
Qy 59 -----LeuTyThrPheAlaProTyTSeValIleIleProGln 71
Db 113 GGAGAGCTTACGATTGAACGCTCAGCTGCTGCGCTGACAGTGGCTTACGGCCTT 172
Qy 72 GLY-----SerTPaSenCyValAlaGluCyValAlaAenThrIleYseSerAp 89
Db 173 GCGCTTCGGCGCGTCTGCTGCTGCTGACGAGATGAAG--AAGACCGCATGTCTATCG 229
Qy 90 LeuGluIlePheGlyArgLeuThr-----ProValGlyAspGlyIleYseMetHisGlu 107
Db 230 CAGGCGCTGTGCACTTCTGCTGCTGCGCGCTGACGAGGAGGAGGAAACGCTGGCC 289
Qy 108 ThrValIleAenGlnMeEPro-----ProGlnAlaAlaVal 119
Db 290 ACTGTGAAGGCGCAAGTCCGCTTCAACAAGTCTGCTGCGCTACCCGCGCGCCGAGC 349
Qy 120 IleLeuLeuThrAspGlyYHIsaenAenLeuGlyMeAenProValGluGluValYseSer 139
Db 350 ATCTCTTGCTGCGAG--CGCTTGAACATGAGATCAACCCCTGTAAGTGTGCTATTC 406
Qy 140 IleTyrgInThrAsnProAenValCySerPheHisValIsePheAlaAspAraIaGlu 159
Db 407 GTCGGCAAG-----AGCGGCGCTGCGCAAGCAACG 436
Qy 160 GlyLyAlaIleIleAspGlnIleValAlaLeuAenSerGlySeValIleuValAspGly 179
Db 437 ATCGCGTCTTATGTCAAGATCATCAAGCTTACGACGCGCTCGTGAACCTGACGCT 496
Qy 180 LeuGlnIleuLeuGlnAenProAlaValCySerGlnGluPheValAenSerValPheYseGln 199
Db 497 GTTACATTGCTACG-----CTGAAGCTTCTGCTGCTTCCC 532
Qy 200 GluGlnIleLeuValThrGluGluValValIleuAraGlyValAspPhe 216
Db 533 AAGCAGATTGGTGTGTCAACACAGAGCATGCTTTTGGCTTCACAT 583

RESULT 30
BH538458/c 699 bp DNA linear GSS 14-DEC-2001
LOCUS BOHRQ84TR BOHR Brassica oleracea genomic clone BOHRQ84, DNA
DEFINITION
sequence.
BHS38458
ACCESSION BHS38458.1 GI:17778057
VERSION
KEYWORDS
SOURCE Brassica oleracea.
ORGANISM Brassica oleracea; Streptophyta; Embryophyta; Tracheophyta;
Eukaryota; Viridiplantae; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE
1 (bases 1 to 699)

```

AUTHORS Town, C.D., Van Aken, S., Utterback, T. and Fraser, C.M.  
TITLE Whole genome shotgun sequencing of Brassica oleracea  
JOURNAL Unpublished (2001)  
COMMENT Other GSSs: BOHRQ84TF  
Contact: Chris Town  
TIGR Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TR  
Class: sheared ends.  
Location/Qualifiers  
1. 699  
/organism="Brassica oleracea"  
/strain="T01000DH3"  
/db\_xref="taxon:3712"  
/clone="BOHRQ84"  
/clone\_1lb="BOHR"  
/notes="Vector: PHOS1; Site 1: BexXI; 2-3 kb sheared  
genomic DNA inserted into PHOS1 using BexXI linkers"

FEATURES  
source  
1. 699  
Location/Qualifiers

BASE COUNT 137 a 170 c 178 g 214 t  
ORIGIN

## Alignment Scores:

Pred. No.:	6.21	Length:	699
Score:	82.00	Matches:	40
Percent Similarity:	33.87%	Conservative:	23
Best Local Similarity:	21.51%	Mismatches:	71
Query Match:	7.33%	Indels:	52
DB:	17	Gaps:	7

US-10-034-500-2 (1-218) x BH538458 (1-699)

```

Qy 31 ValArgGluProLySleGluLeuAlaLyGluAlaIleLeuYslleAaAlaAlaMet 50
Db 426 GTCAACAACAAAAGGCGGAGAGTACGACCGGAGCGTCCACCGCGTAAATAGAGAACTC 367
Qy 51 ProLyMeSeTyrgInGlyGlyLeu-----TyThrPheAla 63
Db 366 GAAGACTCTTGAAGGTGCAATCAATACAGAACTTAGATCCAGACTGCTGCTTAC 307
Qy 64 ProTySeValIleIleProGlnGlySeTyPheSenCyValAlaGluCyValAla 83
Db 306 CCGGTCTGCTCAAAAAGAAATGCGCAATGCGCAGATGTCCAAAGATTC----- 253
Qy 84 AenThrIleYseSerAepLeuGluIlePheGlyArgLeuThrProValGlyAspGlyIle 103
Db 252 -----TTCSCACTCCCGCACATGACCGACTGCTGTAAGCAAGCAAGGGAAC 205
Qy 104 LyseMetHisGluThrValIleAenGlnMeEProProGlnAlaAlaValIleLeuLeuThr 123
Db 204 AAATCTC-----TTATGATTTATGATGCTTCC 178
Qy 124 AepGlyHisAenAenLeuGlyMeAenProValGluGluValYseSerIleTyrgInThr 143
Db 177 TCCGGGCAATATGATATGATATCCGACGATGTAAGAAACGCGTTCATCAC 118
Qy 144 AenProAenVal---CySerPheHisValIsePheAlaAepAraIaGluGlyValAla 162
Db 117 GATCGGGGACATATGCTACAAATATGCTTTCGCT----- 79
Qy 163 IleIleAepGlnIleValAlaLeuAenSerTySeValIleuValAepGlyLeuGlnLeu 182
Db 78 -----CTCAAGAAATGCAAGC----- 64
Qy 183 LeuGlnAenProAlaValCyGlnGluPheValAenSerValPheCyGlnGluGlnIle 202
Db 63 -----GCCACTTACACGACGACTTGTCAATGCAATTTTC---TCCGAACAGATC 19
Qy 203 LeuValThrGluGluVal 208

```







Oy 176 LeuVal-----AspGlyLeuGlnLeu-----LeuGlnAsn 185  
 Db 1472 GCGATCGCTCAGACACTCCATCCAGCTGGAGAGAGGAGTTAATCTCCAGAT 1525  
 RESULT 33  
 LOCUS BE5757248 531 bp mRNA linear EST 25-APR-2001  
 DEFINITION BE5757248 2BOV Bos taurus cDNA 5', mRNA sequence.  
 ACCESSION BE5757248  
 VERSION BE5757248.1 GI:10171240  
 KEYWORDS EST.  
 SOURCE cow.  
 ORGANISM Bos taurus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 Bovidae; Bovinae; Bos.  
 REFERENCE 1 (bases 1 to 531)  
 Smith,T.P.L., Grose,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,  
 Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,  
 G.L., Heaton,M.P., Laegreid,W.M., Rohrer,G.A., Chitko-McKown,C.G.,  
 Perlee,G., Holt,I., Karameyeva,S., Liang,F., Quackenbush,J. and  
 Keeler,J.W.  
 TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA  
 libraries and construction of a gene index for cattle  
 JOURNAL Genome Res. 11 (4), 626-630 (2001)  
 MEDLINE 21180013  
 COMMENT Contact: Smith TPL  
 USDA, ARS, US Meat Animal Research Center  
 PO Box 166, Clay Center, NE 68933-0166, USA  
 Tel: 402 762 4366  
 Fax: 402 762 4390  
 Email: smith@email.marc.usda.gov  
 Single pass sequencing. Bases called and alt trimmed with phred  
 v0.980904.e. Vector identified by cross\_match with the -minscore 18  
 and -mismatch 12 options.  
 PCR Primers  
 FORWARD: AGGAACAGCATGACCAT  
 BACKWARD: GTTCCACGACGACGAG  
 Plate: 63 row: B column: 16  
 Seq primer: ATTAGTGACACTAG.  
 FEATURES  
 source  
 1..531  
 /organism="Bos taurus"  
 /db\_xref="taxon:9913"  
 /clone\_lib="MARC 2BOV"  
 /clone\_type="pooled"  
 /lab\_host="DH10B"  
 /note="Vector: PCMV SPORT6; Site 1: NotI; Site 2: SalI;  
 library made from pooled tissue from testis, thymus,  
 semitendonsus muscle, longissimus muscle, pancreas,  
 adrenal, and endometrium."  
 location/Qualifiers  
 BASE COUNT 139 a 150 c 137 g 105 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 4.57 Length: 531  
 Score: 41.50 Matches: 36  
 Percent Similarity: 44.59% Conservative: 30  
 Best Local Similarity: 24.32% Mismatches: 37  
 Query Match: 7.28% Indels: 45  
 DB: 12 Gaps: 9  
 US-10-034-500-2 (1-218) x BE5757248 (1-531)  
 Oy 16 TyrLeuValAspTyrSerGlySerMetMetLeuHisValAlaValArgGluProLys 35  
 Db 58 TTTGTCATGCACAAGAGCGGCTCATGATGGCAGG-----AAA 96  
 Oy 36 TleGluLeuAlaGlySGuAlaAlaLeuLysAla-----AsnAlaAla 49  
 Db 97 ATAAAGCAGACCGGAGAGGCTTCATCAGATCTGTGATGATCCTCACTCCCATGACGAG 156

Oy 50 MetProLysMetSerTyrGlnGlyLeuTyrThrPheAlaProTyrSerValIleIle 69  
 Db 157 TTGACCTTCATCAGCTTCATGTTCCGAAGCAACACGAGGAAACCT-----TTGCTAGTG 210  
 Oy 70 ProGlnGlySerTyrPheSerCysValAlaLeuLysAlaValAlaThrIleLysSer--- 88  
 Db 211 CCAGCTTCGACTGAGAAC-----GTGAATGAGGCCAAGAGCTAC 249  
 Oy 89 -----AspLeuGluIlePheGlyArgLeuThrProValGlyAspGlyIleLysMetHis 106  
 Db 250 GCCACTGCATTCAGGCCCCAGAGAGG---ACCAATATTAATGATGCATCTGTATGCC 306  
 Oy 107 GluThrValIle-----AsnGlnMetProGlnAlaAla----- 118  
 Db 307 GTGACATCGTGTGAGAAAGCCACAGAGAGAGCTGCTGTAAGGAGCATCACCTC 366  
 Oy 119 ValIleLeuLeuThrAArgGlyHisAsnAsnLeuGlyMetAsnProValGluGluValLys 138  
 Db 367 ATCATCTCTCCATCAGATGAGGAGACCCCATCTGTAGG----- 402  
 Oy 139 SerIleTyrGlnThrAsnProAsn 146  
 Db 403 -----GAGACCAACCTTCG 417  
 RESULT 34  
 LOCUS BF304340 867 bp mRNA linear EST 21-NOV-2000  
 DEFINITION 601887241F1 NIH\_MGC\_17 Homo sapiens cDNA clone IMAGE:4121003 5',  
 mRNA sequence.  
 ACCESSION BF304340  
 VERSION BF304340.1 GI:11251065  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 867)  
 NIH-MGC http://mgc.ncl.nih.gov/.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNU)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNU at: image.lnl.gov  
 Plate: L1CM1002 row: b column: 12  
 High quality sequence stop: 685.  
 FEATURES  
 source  
 1..867  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="IMAGE:4121003"  
 /clone\_lib="NIH\_MGC\_17"  
 /lab\_host="DH10B (phage-resistant)"  
 /tissue\_type="thabdomyosarcoma"  
 /note="Organ: muscle; Vector: pOT7; Site 1: EcoRI;  
 Site 2: XhoI; cDNA made by oligo-dT priming.  
 directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCACGAG(G). Size-selected >500bp  
 for average insert size 1.8kb. Library constructed by  
 Ling Hong in the laboratory of Gerald M. Rubin (University  
 of California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)."  
 location/Qualifiers  
 BASE COUNT 186 a 240 c 245 g 196 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 10.3 Length: 867  
 Score: 81.50 Matches: 39  
 Percent Similarity: 40.54% Conservative: 36



```

AY104471      1713 bp      mRNA      linear      HTC 25-MAY-2002
LOCUS         Zea mays PC0114342 mRNA sequence.
DEFINITION   AY104471
ACCESSION    AY104471
VERSION      AY104471.1 GI:21207549
KEYWORDS     HTC.
SOURCE       Zea mays.
ORGANISM     Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC
              clade; Panicoideae; Andropogoneae; Zea.
REFERENCE    1 (bases 1 to 1713)
AUTHORS      Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitesitt,M.S.,
              Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
              Maize Mapping Project/DuPont Consensus Sequences for Design of
              Overgo Probes
TITLE        Unpublished (2002)
JOURNAL      2 (bases 1 to 1713)
REFERENCE    Coe,E.C.
AUTHORS      Direct Submission
JOURNAL      Submitted (25-Apr-2002) Maize Mapping Project, University of
              Missouri, Columbia, MO 65211, USA
FEATURES     source
              1..1713
              /organism="Zea mays"
              /db_xref="MaizeDB:636659"
              /db_xref="taxon:4577"
              /clone="PC0114342"
              /clone_lib="Maize Mapping Project/DuPont Consensus
              library"
              /note="this sequence is part of a project of EST
              assemblies resulting from the application of public
              contigs to seed DuPont contigs; this resource was
              assembled by DuPont as part of a collaboration for the
              overgo addressing of BACs in conjunction with the Maize
              Mapping Project"
BASE COUNT   433 a 404 c 453 g 443 t
ORIGIN
Alignment Scores:
Pred. No.:   31.8      Length: 1713
Score:        81.50     Matches: 52
Percent Similarity: 32.50% Conservative: 26
Best Local Similarity: 21.67% Mismatches: 55
Query Match:  7.28%     Indels: 107
              Gaps: 13
US-10-034-500-2 (1-218) x AY104471 (1-1713)
QY 48 A1A1A1AmetProlysmetSerTyrgInglyLeuTyThrPhaAlaProTySerVal 67
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 455 GCCCGCTGCCGAGATGAGATGAGAAACGG-----CACCCAGATACCGGGT 505
QY 68 ---11e1leProGInglySerTyPasnSerCysValAlaGlu---CysAlaValAsnThr 85
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 506 GCCGATATTGGACGCGCACTGGCGCAGCGTCCGCCCTCATCGCTCCAAACACC 565
QY 86 11e1ySerAsp-----LeuGlu1lePhelGlyArgLeuThrPro 98
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 566 GCCAAGCTGCTCTCTTATGATGAAGTAAGATGGGTGTGTAAGAAATATGCGCT 625
QY 99 ValGly-----Asp 101
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 626 AACAAGAGAGATATCTGATGCTATTAAAGAACAGATGAGAACTGCAATACTGCCA 685
QY 102 G1Y11e1yMet-----105
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 686 GGATATAAGCTTGAAGAGATGTAATTCGACCTGACTTGAGAGACGAGTCAAGAC 745
QY 106 -----HisGluThrVal-----109
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 746 GCGAATATGCTGTTTGTGACGCCCATCAATTTGTGGAGGGTATATGTAAGAACTT 805

```

```

QY 110 11eAngImetProProGInAla1aVal11eLeuThrAspGly-----125
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 806 GTAGGAGAGTACACACAGAGAGTGAAGCTATCTCCCTCATCAAGGCGATGAGTCAAG 865
QY 126 -----HisAsnLeuGlyMetAsn-----132
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 866 ATGAGAGGCCATGATGATATCCAAAGTTAATCCGGCATACATGGAATATGCTGCT 925
QY 133 -----ProValGluVallySer1le---140
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 926 GTGCTCATGGTCTACATTCGACAAAGAGATGCTGTGCAAGATTGAGGAACACA 985
QY 141 -----TyrgInThrAsnProAsnValCys-----148
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 986 ATTGCGTATAGAAAGATPAGAGAGTGCACAAATCGATGGCTAAACTTTTACACACACC 1045
QY 149 PheHisValAlaSerPhaAlaAspAspAlaGluGlyValAla1le-----1leAsp 165
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1046 TACTTCTAGTTTCTGTGCGCAGAGATATGAAAGAGTAGAGCTGTGTGAACCTTGAAA 1105
QY 166 Gln1leValAlaLeuAsnSerGlySerValLeuValAspGlyLeuGlnLeuGlnAsn 185
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1106 AATATCGTGGCTATTGCGACAGGC-----CTTGATGCTGCTTGATATGGAAACAT 1159

RESULT 37
BG447392      601 bp      mRNA      linear      EST 15-MAR-2001
LOCUS         GA_EB0042101f Gossypium arboreum 7-10 dpa fiber library Gossypium
DEFINITION   GA_EB0042101f Gossypium arboreum 7-10 dpa fiber library Gossypium
ACCESSION    BG447392
VERSION      BG447392.1 GI:13357044
KEYWORDS     EST.
SOURCE       Gossypium arboreum.
ORGANISM     Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
              Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
REFERENCE    1 (bases 1 to 601)
AUTHORS      Wing,R.A., Frischi,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
              D., Wood,T.C., Leslie,A. and Wilkins,T.A.
              An integrated analysis of the genetics, development, and evolution
              of the cotton fiber
              Unpublished (2000)
JOURNAL      Contact: Wing RA
COMMENT      Clemson University Genomics Institute
              Clemson University
              100 Jordan Hall, Clemson, SC 29634, USA
              Tel: 864 656 7286
              Fax: 864 656 4293
              Email: rtwing@clemson.edu
              Seq primer: TAATACGACTCACTATAGG
              High quality sequence stop: 589.
FEATURES     source
              1..601
              /organism="Gossypium arboreum"
              /strain="AKA"
              /cultivar="8400"
              /db_xref="taxon:29729"
              /clone="GA_EB0042101f"
              /clone_lib="Gossypium arboreum 7-10 dpa fiber library"
              /tissue_type="Fibers isolated from bolls harvested 7-10
              dpa"
              /lab_host="E. coli"
              /note="vector: pBR-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT   164 a 114 c 131 g 192 t
ORIGIN
Alignment Scores:
Pred. No.:   6.52      Length: 601
Score:        81.00     Matches: 43
Percent Similarity: 40.38% Conservative: 39
Best Local Similarity: 21.18% Mismatches: 77
Query Match:  7.24%     Indels: 46

```



```

/strain="CC-1690 wild type mt+ 21gr"
/db xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, normalized, lambda zap
II"
/note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:
XhoI; This library, constructed by John Davies and Jeffrey
Medemort, combines cDNAs from CC-1690 cells grown to
mid-log phase in TAP (acetate-containing) medium in the
light, TAP medium in the dark, HS (minimal) medium in
ambient levels of CO2 and HS medium bubbled with 5% CO2.
PolyA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda
zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites.
pBluescript II SK- plasmids were excised from the lambda
ZAP clones by superinfection with Exsist (Stratagene)
phage. The library was normalized using method 4 described
in Bonaldo et al (1996) Genome Research 6: 791-806."
BASE COUNT      119 a      209 c      196 g      132 t
ORIGIN
Alignment Scores:
Pred. No.:      7.54      Length:      656
Score:          81.00      Matches:      44
Percent Similarity: 35.57%      Conservative: 25
Best Local Similarity: 22.68%      Mismatches: 59
Query Match:      7.24%      Indels:      66
DB:              12      Gaps:      11

US-10-034-500-2 (1-218) x BG850892 (1-656)
QY      27 LysHisValAlaValArg-----GluProlyleGluLeuAlaLysGluAlaIle 43
DB      99 CCGCATGTAGTACGCGCGGCCAAGCCGCGCGCGCATGCG--CGTGCAGAGTTTGC 155
QY      44 LeuLysIleAsnAlaIleMetProlyMet-----SeryTyr 55
DB      156 GTCCTCGTGAACGGCGACGCGCCAGATGGCGCAGATTAAAGAGGTGTGCTGCCTAC 215
QY      56 GlnGlyGlyLeuTyrThrPheAlaProTyrSerValIleleProGlnGlySerTrpAn 75
DB      216 TCGGGGGGCGCTGGACACC-----TTCATCATCTCTC-----AAGTGGCTG 254
QY      76 SerCyValAlaGluCyValAlaValAsnThrIleLysSerArgLeuIlePheGlyArg 95
DB      255 CAGGACACCTACGCGCTGCGAGGTGTCACCTTACCGCGCACTG----- 239
QY      96 LeuThrProValGlyAspGlyIleLysMetHisGluThrValIleAsnGlnMetProPro 115
DB      300 -----GGCCAGGCGCGAGAGCTGGAG-----CCTGCC 326
QY      116 GlnAlaIleValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsnProValGlu 135
DB      327 CGCGCCAGAGGCC-----GAGAAAGATGGGC----- 350
QY      136 GluValLysSerIleTyrGlnThrAsnProAsnValCyAspHeHisValIleSerPheAla 155
DB      351 ---GTCAGAGAGTC-----TTTATT 368
QY      156 AspAspAlaGlnGlyLysAlaIleIleAspGlnIleValAlaLeuAsnSerGlySerVal 175
DB      369 GAGGACCTGCGCGAGAGGATTGTCGCGGACTACGTTTCCCATGTCAGGGCAACGCG 428
QY      176 LeuValAspGlyLeuGlnLeuLeuGlnAsnProAlaValCySerGlnGluPheValAsnSer 195
DB      429 CTATTACGAGGGCACTTACCTGCTGGCACC-----TTCATTGCGCGCCCG 473
QY      196 ValPheCySerGlnGlnGlnIleLeuValThrGlnGluValVal 209
DB      474 CTCATGCGCCAGGCGCCAGATCGAGATCGCCAGAGAGTGTC 515

RESULT 40
A1684487/c      764 bp      mRNA      linear      EST 07-MAR-2000
LOCUS      A1684487

```

```

DEFINITION      wa82f03.x1 Soares NFL T GBC S1 Homo sapiens cDNA clone
IMAGE:2302685.3 similar to TR:075602 075602 SPERM FLAGELLAR
PROTEIN REPO-SA-1.; mRNA sequence.
ACCESSION      A1684487.1 GI:4895781
VERSION        A1684487.1
KEYWORDS
SOURCE
ORGANISM      human.
               Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS        NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE          National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
               Tumor Gene Index
JOURNAL        Unpublished (1997)
COMMENT        Contact: Robert Strausberg, Ph.D.
               Email: cgaps-remail.nih.gov
               This clone is available royalty-free through LNL; contact the
               IMAGE Consortium (info@image.lnl.gov) for further information.
               Insert Length: 1814 Std Error: 0.00
               Seq primer: -40UP from G1pco
               High quality sequence stop: 461.
               Location/Qualifiers
               1..764
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone="IMAGE:2302685"
               /clone_lib="Soares_NFL_T_GBC_S1"
               /lab_host="DH10B"
               /note="Organ: pooled; Vector: pTTT3D-Pac (Pharmacia) with
               a modified polylinker; Site 1: Not I; Site 2: Eco RI;
               Equal amounts of plasmid DNA from three normalized
               libraries (fetal lung NDHL19W, testis NHT, and B-cell
               NCI CGAP GCBI) were mixed, and as circles were made in
               vitro. Following HAP purification, this DNA was used as
               tracer in a subtractive hybridization reaction. The driver
               was PCR-amplified cDNAs from pools of 5,000 clones made
               from the same 3 libraries. The pools consisted of
               1 M.A.G.E. clones 297480-302087, 682632-687239,
               726408-728711, and 729096-731399. Subtraction by Benito
               Soares and M. Patricia Bonaldo."
BASE COUNT      205 a      163 c      163 g      233 t
ORIGIN
Alignment Scores:
Pred. No.:      9.71      Length:      764
Score:          81.00      Matches:      45
Percent Similarity: 37.02%      Conservative: 32
Best Local Similarity: 21.63%      Mismatches: 55
Query Match:      7.24%      Indels:      76
DB:              9      Gaps:      10

US-10-034-500-2 (1-218) x A1684487 (1-764)
QY      22 GlySerMetMetLysHisValAlaValArgGluProlyleGluLeuAlaLysGlu 41
DB      670 GGCATCATGATGCTGTTATGTAGAGAGCTCATCTGAGACCATGACATGCA----- 617
QY      42 AlaIleLeuLysIleAsnAlaIleMetProlyMetSeryTyrGlnGlyLeuTyrThr 61
DB      617 ----- 617
QY      62 PheAlaProTyrSerValIleleProGlnGly-----SeryPAsnSerCyValAla 79
DB      616 -----GTCATCATTTCTTAAGG5GTAAACCAAGTTTGTCACTGTGCTTGTCA 572
QY      80 GluCyAlaValAsnThrIleLysSer-----AspLeuGlnIlePheGlyArg 95
DB      571 GAAGAACCGGAGAGATATTTAAGCTGACAGCTGTGGCCTTATGACAGATTGGAGA 512
QY      96 LeuThrProValGlyAspGlyIleLysMetHisGlu-----ThrValIleAsnGln 112
DB      511 CACACTCTCT-----GAACACGACGCGGCTGTGTGCACTCAACAATACT 470

```







GenCore version 5.1.4 p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 9, 2003, 14:41:28 ; Search time 28 seconds  
(without alignments)  
229.078 Million cell updates/sec

Title: US-10-034-500-2

Perfect score: 218  
Sequence: 1 AEVTASCTKRVESYNVLVDY.....OEQILVTEVVVLGKVFAP 218

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size: 0

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database:

Issued Patents AA: \*  
1: /cgn2\_6/ptodata/1/iaa/5A.COMB.pep: \*  
2: /cgn2\_6/ptodata/1/iaa/5B.COMB.pep: \*  
3: /cgn2\_6/ptodata/1/iaa/6A.COMB.pep: \*  
4: /cgn2\_6/ptodata/1/iaa/6B.COMB.pep: \*  
5: /cgn2\_6/ptodata/1/iaa/PCTus.COMB.pep: \*  
6: /cgn2\_6/ptodata/1/iaa/backfilest.pep: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	3.2	209	4	US-09-196-293-15
2	7	3.2	209	4	US-08-235-836C-32
3	7	3.2	210	1	US-08-158-353-3
4	7	3.2	210	4	US-08-209-603E-15
5	7	3.2	210	4	US-08-235-836C-30
6	7	3.2	238	4	US-09-370-838-37
7	7	3.2	243	4	US-09-370-838-39
8	7	3.2	243	4	US-09-370-838-42
9	7	3.2	244	4	US-09-370-838-43
10	7	3.2	244	4	US-09-370-838-46
11	7	3.2	245	4	US-09-370-838-40
12	7	3.2	466	4	US-08-235-836C-107
13	7	3.2	466	4	US-08-235-836C-110
14	7	3.2	588	4	US-08-235-836C-122
15	7	3.2	749	4	US-09-562-737-95
16	7	3.2	858	3	US-08-946-026-3
17	7	3.2	1005	4	US-09-206-942-41
18	7	3.2	1011	4	US-09-206-942-39
19	6	2.8	107	4	US-09-162-564-6
20	6	2.8	121	4	US-08-207-481-31
21	6	2.8	121	5	PCT-US95-02689-33
22	6	2.8	153	3	US-08-993-176-8
23	6	2.8	178	4	US-09-040-213-2
24	6	2.8	202	4	US-09-026-958-2
25	6	2.8	202	4	US-09-057-860A-2
26	6	2.8	207	2	US-08-177-109A-60
27	6	2.8	207	2	US-08-687-706-60

28	6	2.8	212	1	US-08-031-295-2	Sequence 2, Appli
29	6	2.8	212	4	US-07-903-580-2	Sequence 2, Appli
30	6	2.8	226	4	US-09-134-001C-2945	Sequence 2945, Ap
31	6	2.8	242	2	US-08-512-955-4	Sequence 4, Appli
32	6	2.8	246	4	US-09-627-376-14	Sequence 14, Appli
33	6	2.8	246	4	US-09-452-239-20	Sequence 20, Appli
34	6	2.8	246	4	US-09-452-239-22	Sequence 22, Appli
35	6	2.8	247	4	US-09-452-239-18	Sequence 18, Appli
36	6	2.8	247	4	US-09-452-239-10	Sequence 50, Appli
37	6	2.8	248	4	US-09-189-060B-12	Sequence 12, Appli
38	6	2.8	248	4	US-09-452-239-24	Sequence 24, Appli
39	6	2.8	248	4	US-09-452-239-26	Sequence 26, Appli
40	6	2.8	254	2	US-08-207-481-20	Sequence 20, Appli
41	6	2.8	254	5	PCT-US95-02689-20	Sequence 20, Appli
42	6	2.8	259	4	US-09-452-239-42	Sequence 42, Appli
43	6	2.8	260	4	US-09-452-239-12	Sequence 12, Appli
44	6	2.8	264	2	US-08-484-905-120	Sequence 120, App
45	6	2.8	264	3	US-08-481-985B-120	Sequence 120, App
46	6	2.8	264	4	US-08-370-476-120	Sequence 120, App
47	6	2.8	271	3	US-08-968-563-33	Sequence 33, Appli
48	6	2.8	271	4	US-08-969-683A-33	Sequence 33, Appli
49	6	2.8	271	4	US-09-297-928-14	Sequence 14, Appli
50	6	2.8	276	1	US-08-431-387-7	Sequence 7, Appli
51	6	2.8	282	1	US-08-446-924-4	Sequence 4, Appli
52	6	2.8	282	1	US-08-798-665-4	Sequence 4, Appli
53	6	2.8	282	3	US-08-982-987A-4	Sequence 4, Appli
54	6	2.8	312	4	US-09-267-177-21	Sequence 21, Appli
55	6	2.8	320	3	US-09-217-609A-6	Sequence 6, Appli
56	6	2.8	320	4	US-08-873-255B-6	Sequence 6, Appli
57	6	2.8	322	4	US-09-424-349A-4	Sequence 4, Appli
58	6	2.8	344	1	US-08-400-422-4	Sequence 4, Appli
59	6	2.8	357	4	US-08-838-151A-20	Sequence 20, Appli
60	6	2.8	357	4	US-08-838-151A-24	Sequence 24, Appli
61	6	2.8	357	4	US-08-838-151A-27	Sequence 27, Appli
62	6	2.8	357	4	US-08-838-151A-30	Sequence 30, Appli
63	6	2.8	365	2	US-08-827-190-5	Sequence 5, Appli
64	6	2.8	365	3	US-09-231-529-3	Sequence 3, Appli
65	6	2.8	365	4	US-08-977-816-3	Sequence 3, Appli
66	6	2.8	365	4	US-09-170-187-5	Sequence 5, Appli
67	6	2.8	369	2	US-08-723-415B-4	Sequence 4, Appli
68	6	2.8	369	4	US-09-189-627A-4	Sequence 4, Appli
69	6	2.8	369	4	US-09-710-861-4	Sequence 4, Appli
70	6	2.8	370	2	US-08-723-415B-6	Sequence 6, Appli
71	6	2.8	370	4	US-09-189-627A-6	Sequence 6, Appli
72	6	2.8	370	4	US-09-710-861-6	Sequence 6, Appli
73	6	2.8	385	2	US-08-723-415B-8	Sequence 8, Appli
74	6	2.8	385	4	US-09-189-627A-8	Sequence 8, Appli
75	6	2.8	385	4	US-09-710-861-8	Sequence 8, Appli
76	6	2.8	387	4	US-09-134-001C-4298	Sequence 4298, Ap
77	6	2.8	418	4	US-09-198-603C-10	Sequence 30, Appli
78	6	2.8	422	3	US-09-109-204-30	Sequence 30, Appli
79	6	2.8	424	4	US-09-134-001C-4950	Sequence 4950, Ap
80	6	2.8	431	4	US-09-071-035-44	Sequence 344, App
81	6	2.8	432	3	US-08-985-908-31	Sequence 31, Appli
82	6	2.8	438	4	US-09-404-879A-390	Sequence 390, App
83	6	2.8	446	2	US-08-723-415B-2	Sequence 2, Appli
84	6	2.8	446	4	US-09-189-627A-2	Sequence 2, Appli
85	6	2.8	446	4	US-09-189-627A-267	Sequence 267, App
86	6	2.8	446	4	US-09-710-861-2	Sequence 2, Appli
87	6	2.8	447	3	US-08-508-761B-6	Sequence 6, Appli
88	6	2.8	456	4	US-08-840-767-11	Sequence 11, Appli
89	6	2.8	464	4	US-08-580-031A-14	Sequence 14, Appli
90	6	2.8	465	3	US-08-701-582D-13	Sequence 13, Appli
91	6	2.8	465	3	US-09-063-893A-19	Sequence 19, Appli
92	6	2.8	465	4	US-08-840-767-8	Sequence 8, Appli
93	6	2.8	465	4	US-08-840-767-48	Sequence 48, Appli
94	6	2.8	465	4	US-09-096-776B-6	Sequence 6, Appli
95	6	2.8	474	4	US-09-096-776B-10	Sequence 10, Appli
96	6	2.8	474	4	US-09-071-035-142	Sequence 342, App
97	6	2.8	475	4	US-08-840-767-10	Sequence 10, Appli
98	6	2.8	492	4	US-09-342-749-2	Sequence 2, Appli
99	6	2.8	492	4	US-09-691-940-2	Sequence 2, Appli
100	6	2.8	494	4	US-09-019-095A-24	Sequence 24, Appli

101	6	2.8	497	4	US-09-058-947A-4	Sequence 4, Appli	174	5	2.3	12	5	PCT-US95-04886-8	Sequence 8, Appli
102	6	2.8	500	4	US-08-868-373-12	Sequence 12, Appli	175	5	2.3	13	1	US-08-346-455B-32	Sequence 32, Appli
103	6	2.8	506	3	US-08-888-998-2	Sequence 2, Appli	176	5	2.3	13	3	US-08-977-221-32	Sequence 32, Appli
104	6	2.8	506	4	US-09-362-633-2	Sequence 2, Appli	177	5	2.3	13	4	US-09-483-831B-32	Sequence 32, Appli
105	6	2.8	537	4	US-08-868-373-10	Sequence 10, Appli	178	5	2.3	13	5	PCT-US95-06613-32	Sequence 32, Appli
106	6	2.8	563	4	US-09-718-841-2	Sequence 2, Appli	179	5	2.3	15	5	PCT-US93-06751-55	Sequence 55, Appli
107	6	2.8	563	4	US-09-718-841-2	Sequence 2, Appli	180	5	2.3	16	1	US-07-752-101A-15	Sequence 15, Appli
108	6	2.8	596	2	US-08-752-307B-13	Sequence 13, Appli	181	5	2.3	16	2	US-08-591-438-10	Sequence 10, Appli
109	6	2.8	596	2	US-09-707-802-13	Sequence 13, Appli	182	5	2.3	16	4	US-09-323-433A-18	Sequence 18, Appli
110	6	2.8	596	4	US-09-991-326-13	Sequence 13, Appli	183	5	2.3	16	4	US-09-293-322C-3	Sequence 3, Appli
111	6	2.8	645	4	US-09-202-712-2	Sequence 2, Appli	184	5	2.3	16	6	US-09-293-322C-3	Sequence 3, Appli
112	6	2.8	645	4	US-08-630-916A-46	Sequence 46, Appli	185	5	2.3	17	2	US-08-719-758-7	Sequence 7, Appli
113	6	2.8	743	3	US-09-738-884-4	Sequence 4, Appli	186	5	2.3	17	2	US-08-729-152-35	Sequence 35, Appli
114	6	2.8	744	4	US-09-738-884-5	Sequence 5, Appli	187	5	2.3	17	4	US-08-974-549A-67	Sequence 67, Appli
115	6	2.8	749	4	US-09-562-737-91	Sequence 91, Appli	188	5	2.3	17	4	US-09-119-827-7	Sequence 7, Appli
116	6	2.8	749	4	US-09-562-737-92	Sequence 92, Appli	189	5	2.3	18	1	US-08-338-127-2	Sequence 2, Appli
117	6	2.8	749	4	US-09-562-737-93	Sequence 93, Appli	190	5	2.3	18	2	US-09-017-205-19	Sequence 19, Appli
118	6	2.8	749	4	US-09-562-737-94	Sequence 94, Appli	191	5	2.3	19	4	US-08-974-549A-636	Sequence 636, App
119	6	2.8	795	4	US-09-193-562D-11	Sequence 11, Appli	192	5	2.3	20	1	US-08-338-127-1	Sequence 1, Appli
120	6	2.8	805	1	US-08-045-806-2	Sequence 2, Appli	193	5	2.3	20	1	US-08-484-135-59	Sequence 59, Appli
121	6	2.8	805	1	US-08-366-051B-2	Sequence 2, Appli	194	5	2.3	20	1	US-08-484-135-60	Sequence 60, Appli
122	6	2.8	816	2	US-08-533-306A-6	Sequence 6, Appli	195	5	2.3	20	1	US-08-484-135-62	Sequence 92, Appli
123	6	2.8	816	2	US-08-742-923A-6	Sequence 6, Appli	196	5	2.3	20	1	US-08-484-135-63	Sequence 93, Appli
124	6	2.8	831	4	US-09-193-562D-12	Sequence 12, Appli	197	5	2.3	20	1	US-08-484-635-59	Sequence 59, Appli
125	6	2.8	831	4	US-09-404-879A-389	Sequence 389, App	198	5	2.3	20	1	US-08-484-635-60	Sequence 60, Appli
126	6	2.8	852	2	US-09-070-060-3	Sequence 3, Appli	199	5	2.3	20	1	US-08-484-635-61	Sequence 61, Appli
127	6	2.8	852	2	US-09-357-746-3	Sequence 3, Appli	200	5	2.3	20	2	US-08-484-631-59	Sequence 59, Appli
128	6	2.8	854	2	US-09-070-060-4	Sequence 4, Appli	201	5	2.3	20	2	US-08-484-631-60	Sequence 60, Appli
129	6	2.8	854	2	US-09-357-746-4	Sequence 4, Appli	202	5	2.3	20	2	US-08-484-631-61	Sequence 61, Appli
130	6	2.8	874	3	US-08-804-433A-15	Sequence 15, Appli	203	5	2.3	20	2	US-08-827-570-59	Sequence 59, Appli
131	6	2.8	874	3	US-08-720-229-15	Sequence 15, Appli	204	5	2.3	20	2	US-08-827-570-60	Sequence 60, Appli
132	6	2.8	885	2	US-08-533-306A-4	Sequence 4, Appli	205	5	2.3	20	2	US-08-827-570-61	Sequence 61, Appli
133	6	2.8	885	2	US-08-742-923A-4	Sequence 4, Appli	206	5	2.3	21	2	US-08-789-078-6	Sequence 6, Appli
134	6	2.8	905	4	US-09-193-562D-46	Sequence 46, Appli	207	5	2.3	21	2	US-08-789-078-6	Sequence 6, Appli
135	6	2.8	905	4	US-09-193-562D-2	Sequence 2, Appli	208	5	2.3	21	2	US-08-789-078-6	Sequence 6, Appli
136	6	2.8	914	4	US-09-404-879A-312	Sequence 312, App	209	5	2.3	21	2	US-08-789-078-6	Sequence 6, Appli
137	6	2.8	914	4	US-09-139-802-201	Sequence 201, App	210	5	2.3	21	3	US-08-789-078-6	Sequence 6, Appli
138	6	2.8	1000	4	US-09-193-562D-30	Sequence 30, Appli	211	5	2.3	21	5	PCT-US95-04886-4	Sequence 4, Appli
139	6	2.8	1119	4	US-09-396-651B-2	Sequence 2, Appli	212	5	2.3	21	5	PCT-US95-04886-6	Sequence 6, Appli
140	6	2.8	1151	4	US-09-134-001C-3242	Sequence 3242, Ap	213	5	2.3	22	1	US-07-933-469A-5	Sequence 5, Appli
141	6	2.8	1170	4	US-09-749-588-2	Sequence 2, Appli	214	5	2.3	22	1	US-07-933-469A-6	Sequence 6, Appli
142	6	2.8	1337	3	US-08-854-585-2	Sequence 2, Appli	215	5	2.3	22	1	US-07-752-101A-5	Sequence 5, Appli
143	6	2.8	1337	5	PCT-US95-05512-2	Sequence 2, Appli	216	5	2.3	22	1	US-08-250-310-5	Sequence 5, Appli
144	6	2.8	1408	1	US-08-612-521-2	Sequence 2, Appli	217	5	2.3	22	1	US-08-250-310-6	Sequence 6, Appli
145	6	2.8	1788	2	US-08-962-284-2	Sequence 2, Appli	218	5	2.3	22	1	US-08-439-404-5	Sequence 5, Appli
146	6	2.8	1792	2	US-08-962-284-4	Sequence 4, Appli	219	5	2.3	22	1	US-08-439-404-6	Sequence 6, Appli
147	6	2.8	3567	2	US-07-642-734C-4	Sequence 4, Appli	220	5	2.3	22	2	US-08-789-078-5	Sequence 5, Appli
148	6	2.8	3567	2	US-08-439-009A-4	Sequence 4, Appli	221	5	2.3	22	2	US-08-789-078-5	Sequence 5, Appli
149	5	2.3	3567	4	US-08-974-549A-21	Sequence 21, Appli	222	5	2.3	22	3	US-08-833-327-5	Sequence 3, Appli
150	5	2.3	7	4	US-08-974-549A-130	Sequence 130, App	223	5	2.3	22	3	US-08-833-327-6	Sequence 6, Appli
151	5	2.3	8	1	US-08-267-092A-23	Sequence 23, Appli	224	5	2.3	22	3	US-09-340-781B-5	Sequence 5, Appli
152	5	2.3	8	2	US-08-540-412-51	Sequence 51, Appli	225	5	2.3	22	3	US-09-340-781B-6	Sequence 6, Appli
153	5	2.3	8	2	US-08-926-412-51	Sequence 51, Appli	226	5	2.3	22	4	US-08-974-549A-84	Sequence 84, Appli
154	5	2.3	8	3	US-09-112-656-60	Sequence 60, Appli	227	5	2.3	22	4	US-08-974-549A-84	Sequence 84, Appli
155	5	2.3	8	4	US-09-051-342-23	Sequence 23, Appli	228	5	2.3	22	5	PCT-US95-04886-1	Sequence 1, Appli
156	5	2.3	8	4	US-08-468-161-23	Sequence 23, Appli	229	5	2.3	24	1	US-07-752-101A-5	Sequence 5, Appli
157	5	2.3	8	4	US-09-051-759-23	Sequence 23, Appli	230	5	2.3	24	1	US-08-103-170-26	Sequence 26, Appli
158	5	2.3	8	4	US-09-254-892-51	Sequence 51, Appli	231	5	2.3	24	2	US-08-103-170-27	Sequence 27, Appli
159	5	2.3	8	5	PCT-US95-08156-23	Sequence 23, Appli	232	5	2.3	24	2	US-08-651-420A-22	Sequence 22, Appli
160	5	2.3	10	3	US-08-396-385-3	Sequence 3, Appli	233	5	2.3	24	4	US-09-157-864-6	Sequence 6, Appli
161	5	2.3	10	3	US-08-159-339A-958	Sequence 958, App	234	5	2.3	25	1	US-08-338-127-3	Sequence 3, Appli
162	5	2.3	10	4	US-08-159-339A-1064	Sequence 1064, Ap	235	5	2.3	25	2	US-08-473-475A-5	Sequence 5, Appli
163	5	2.3	10	4	US-09-287-221-3	Sequence 3, Appli	236	5	2.3	25	2	US-08-637-759B-332	Sequence 332, App
164	5	2.3	11	2	US-08-789-078-9	Sequence 9, Appli	237	5	2.3	25	3	US-08-871-355A-232	Sequence 232, App
165	5	2.3	11	2	US-08-752-633-9	Sequence 9, Appli	238	5	2.3	25	4	US-09-028-937-42	Sequence 42, Appli
166	5	2.3	11	5	PCT-US95-04886-9	Sequence 9, Appli	239	5	2.3	25	4	US-09-028-937-43	Sequence 43, Appli
167	5	2.3	12	2	US-08-789-078-7	Sequence 7, Appli	240	5	2.3	25	4	US-09-028-937-44	Sequence 44, Appli
168	5	2.3	12	2	US-08-789-078-8	Sequence 8, Appli	241	5	2.3	25	4	US-09-201-945-232	Sequence 232, App
169	5	2.3	12	2	US-08-752-633-7	Sequence 7, Appli	242	5	2.3	25	4	US-09-288-143-94	Sequence 94, Appli
170	5	2.3	12	2	US-08-752-633-8	Sequence 8, Appli	243	5	2.3	26	2	US-08-620-151-121	Sequence 121, App
171	5	2.3	12	3	US-08-844-978-7	Sequence 4, Appli	244	5	2.3	27	3	US-08-851-843A-157	Sequence 157, App
172	5	2.3	12	3	US-08-844-978-7	Sequence 7, Appli	245	5	2.3	27	3	US-08-851-843A-169	Sequence 169, App
173	5	2.3	12	5	PCT-US95-04886-7	Sequence 7, Appli	246	5	2.3	27	4	US-08-974-549A-877	Sequence 277, App

247	5	2.3	27	4	US-08-974-549A-289	Sequence 269, App	320	5	2.3	35	3	US-08-484-223B-36	Sequence 36, App1
248	5	2.3	27	4	US-08-854-050-157	Sequence 167, App	321	5	2.3	35	3	US-08-484-223B-37	Sequence 37, App1
249	5	2.3	27	4	US-08-854-050-169	Sequence 167, App	322	5	2.3	35	3	US-08-484-223B-38	Sequence 38, App1
250	5	2.3	27	4	US-09-430-323-157	Sequence 157, App	323	5	2.3	35	3	US-08-484-223B-39	Sequence 39, App1
251	5	2.3	27	4	US-09-430-323-169	Sequence 159, App	324	5	2.3	35	3	US-08-484-223B-40	Sequence 40, App1
252	5	2.3	28	4	US-09-040-485-3	Sequence 3, App11	325	5	2.3	35	3	US-08-484-223B-41	Sequence 41, App1
253	5	2.3	28	4	US-09-252-586-17	Sequence 11, App1	326	5	2.3	35	3	US-08-484-223B-42	Sequence 42, App1
254	5	2.3	30	1	US-08-244-701B-25	Sequence 25, App1	327	5	2.3	35	3	US-08-484-223B-43	Sequence 43, App1
255	5	2.3	30	1	US-08-244-701B-28	Sequence 28, App1	328	5	2.3	35	3	US-08-484-223B-44	Sequence 44, App1
256	5	2.3	30	2	US-08-846-762-58	Sequence 58, App1	329	5	2.3	35	3	US-08-484-223B-45	Sequence 45, App1
257	5	2.3	30	4	US-08-818-112-62	Sequence 62, App1	330	5	2.3	35	3	US-08-484-223B-46	Sequence 46, App1
258	5	2.3	30	4	US-08-818-111-62	Sequence 62, App1	331	5	2.3	35	3	US-08-484-223B-47	Sequence 47, App1
259	5	2.3	30	4	US-09-056-556-62	Sequence 25, App1	332	5	2.3	35	3	US-08-484-223B-48	Sequence 48, App1
260	5	2.3	30	4	US-09-076-721-25	Sequence 25, App1	333	5	2.3	35	3	US-08-484-223B-49	Sequence 49, App1
261	5	2.3	30	4	US-09-076-721-28	Sequence 28, App1	334	5	2.3	35	3	US-08-484-223B-50	Sequence 50, App1
262	5	2.3	30	4	US-09-072-596-62	Sequence 62, App1	335	5	2.3	35	3	US-08-484-223B-51	Sequence 51, App1
263	5	2.3	32	1	US-08-471-780C-129	Sequence 129, App	336	5	2.3	35	3	US-08-484-223B-52	Sequence 52, App1
264	5	2.3	32	1	US-08-467-282B-129	Sequence 129, App	337	5	2.3	35	3	US-08-484-223B-53	Sequence 53, App1
265	5	2.3	32	2	US-08-471-282A-129	Sequence 129, App	338	5	2.3	35	3	US-08-484-223B-54	Sequence 54, App1
266	5	2.3	32	2	US-08-466-710C-129	Sequence 129, App	339	5	2.3	35	3	US-08-484-223B-55	Sequence 55, App1
267	5	2.3	32	3	US-08-468-739C-129	Sequence 129, App	340	5	2.3	35	3	US-08-484-223B-56	Sequence 56, App1
268	5	2.3	32	3	US-08-851-843A-136	Sequence 136, App	341	5	2.3	35	3	US-08-484-223B-57	Sequence 57, App1
269	5	2.3	32	4	US-08-974-549A-254	Sequence 254, App	342	5	2.3	35	3	US-08-484-223B-58	Sequence 58, App1
270	5	2.3	32	4	US-08-854-050-136	Sequence 136, App	343	5	2.3	35	3	US-08-484-223B-59	Sequence 59, App1
271	5	2.3	32	4	US-09-220-528-117	Sequence 117, App	344	5	2.3	35	3	US-08-484-223B-60	Sequence 60, App1
272	5	2.3	32	4	US-09-430-323-136	Sequence 136, App	345	5	2.3	35	3	US-08-484-223B-61	Sequence 61, App1
273	5	2.3	32	4	US-09-314-268-122	Sequence 122, App	346	5	2.3	35	3	US-08-484-223B-62	Sequence 62, App1
274	5	2.3	32	4	US-09-131-750-17	Sequence 17, App1	347	5	2.3	35	3	US-08-484-223B-63	Sequence 63, App1
275	5	2.3	33	1	US-08-244-701B-1	Sequence 1, App11	348	5	2.3	35	3	US-08-484-223B-64	Sequence 64, App1
276	5	2.3	33	4	US-09-076-721-1	Sequence 1, App11	349	5	2.3	35	3	US-08-484-223B-65	Sequence 65, App1
277	5	2.3	34	4	US-09-082-279B-381	Sequence 381, App	350	5	2.3	35	3	US-08-484-223B-66	Sequence 66, App1
278	5	2.3	34	4	US-09-315-304B-381	Sequence 381, App	351	5	2.3	35	3	US-08-484-223B-67	Sequence 67, App1
279	5	2.3	35	2	US-08-700-442A-20	Sequence 20, App1	352	5	2.3	35	3	US-08-484-223B-68	Sequence 68, App1
280	5	2.3	35	2	US-08-651-420A-23	Sequence 23, App1	353	5	2.3	35	3	US-08-484-223B-69	Sequence 69, App1
281	5	2.3	35	3	US-08-486-099-18	Sequence 18, App1	354	5	2.3	35	3	US-08-484-223B-70	Sequence 70, App1
282	5	2.3	35	3	US-08-486-099-53	Sequence 33, App1	355	5	2.3	35	3	US-08-484-223B-71	Sequence 71, App1
283	5	2.3	35	3	US-08-486-099-34	Sequence 34, App1	356	5	2.3	35	3	US-08-484-223B-72	Sequence 72, App1
284	5	2.3	35	3	US-08-486-099-35	Sequence 35, App1	357	5	2.3	35	3	US-08-484-223B-73	Sequence 73, App1
285	5	2.3	35	3	US-08-486-099-36	Sequence 36, App1	358	5	2.3	35	3	US-08-484-223B-74	Sequence 74, App1
286	5	2.3	35	3	US-08-486-099-37	Sequence 37, App1	359	5	2.3	35	3	US-08-484-223B-75	Sequence 75, App1
287	5	2.3	35	3	US-08-486-099-38	Sequence 38, App1	360	5	2.3	35	3	US-08-484-223B-76	Sequence 76, App1
288	5	2.3	35	3	US-08-486-099-39	Sequence 39, App1	361	5	2.3	35	3	US-08-484-223B-77	Sequence 77, App1
289	5	2.3	35	3	US-08-486-099-40	Sequence 40, App1	362	5	2.3	35	3	US-08-484-223B-78	Sequence 78, App1
290	5	2.3	35	3	US-08-486-099-41	Sequence 41, App1	363	5	2.3	35	3	US-08-484-223B-79	Sequence 79, App1
291	5	2.3	35	3	US-08-486-099-42	Sequence 42, App1	364	5	2.3	35	3	US-08-484-223B-80	Sequence 80, App1
292	5	2.3	35	3	US-08-486-099-43	Sequence 43, App1	365	5	2.3	35	3	US-08-484-223B-81	Sequence 81, App1
293	5	2.3	35	3	US-08-486-099-44	Sequence 44, App1	366	5	2.3	35	3	US-08-484-223B-82	Sequence 82, App1
294	5	2.3	35	3	US-08-486-099-45	Sequence 45, App1	367	5	2.3	35	3	US-08-484-223B-83	Sequence 83, App1
295	5	2.3	35	3	US-08-486-099-46	Sequence 46, App1	368	5	2.3	35	3	US-08-484-223B-84	Sequence 84, App1
296	5	2.3	35	3	US-08-486-099-47	Sequence 47, App1	369	5	2.3	35	3	US-08-484-223B-85	Sequence 85, App1
297	5	2.3	35	3	US-08-486-099-48	Sequence 48, App1	370	5	2.3	35	3	US-08-484-223B-86	Sequence 86, App1
298	5	2.3	35	3	US-08-486-099-49	Sequence 49, App1	371	5	2.3	35	3	US-08-484-223B-87	Sequence 87, App1
299	5	2.3	35	3	US-08-360-107A-153	Sequence 153, App	372	5	2.3	35	3	US-08-484-223B-88	Sequence 88, App1
300	5	2.3	35	3	US-08-360-107A-18	Sequence 18, App1	373	5	2.3	35	3	US-08-484-223B-89	Sequence 89, App1
301	5	2.3	35	3	US-08-360-107A-33	Sequence 33, App1	374	5	2.3	35	3	US-08-484-223B-90	Sequence 90, App1
302	5	2.3	35	3	US-08-360-107A-34	Sequence 34, App1	375	5	2.3	35	3	US-08-484-223B-91	Sequence 91, App1
303	5	2.3	35	3	US-08-360-107A-35	Sequence 35, App1	376	5	2.3	35	3	US-08-484-223B-92	Sequence 92, App1
304	5	2.3	35	3	US-08-360-107A-36	Sequence 36, App1	377	5	2.3	35	3	US-08-484-223B-93	Sequence 93, App1
305	5	2.3	35	3	US-08-360-107A-37	Sequence 37, App1	378	5	2.3	35	3	US-08-484-223B-94	Sequence 94, App1
306	5	2.3	35	3	US-08-360-107A-38	Sequence 38, App1	379	5	2.3	35	3	US-08-484-223B-95	Sequence 95, App1
307	5	2.3	35	3	US-08-360-107A-39	Sequence 39, App1	380	5	2.3	35	3	US-08-484-223B-96	Sequence 96, App1
308	5	2.3	35	3	US-08-360-107A-40	Sequence 40, App1	381	5	2.3	35	3	US-08-484-223B-97	Sequence 97, App1
309	5	2.3	35	3	US-08-360-107A-41	Sequence 41, App1	382	5	2.3	35	3	US-08-484-223B-98	Sequence 98, App1
310	5	2.3	35	3	US-08-360-107A-42	Sequence 42, App1	383	5	2.3	35	3	US-08-484-223B-99	Sequence 99, App1
311	5	2.3	35	3	US-08-360-107A-43	Sequence 43, App1	384	5	2.3	35	3	US-08-484-223B-100	Sequence 100, App1
312	5	2.3	35	3	US-08-360-107A-44	Sequence 44, App1	385	5	2.3	35	3	US-08-484-223B-101	Sequence 101, App1
313	5	2.3	35	3	US-08-360-107A-45	Sequence 45, App1	386	5	2.3	35	3	US-08-484-223B-102	Sequence 102, App1
314	5	2.3	35	3	US-08-360-107A-46	Sequence 46, App1	387	5	2.3	35	3	US-08-484-223B-103	Sequence 103, App1
315	5	2.3	35	3	US-08-360-107A-47	Sequence 47, App1	388	5	2.3	35	3	US-08-484-223B-104	Sequence 104, App1
316	5	2.3	35	3	US-08-360-107A-48	Sequence 48, App1	389	5	2.3	35	3	US-08-484-223B-105	Sequence 105, App1
317	5	2.3	35	3	US-08-484-223B-18	Sequence 18, App1	390	5	2.3	35	3	US-08-484-223B-106	Sequence 106, App1
318	5	2.3	35	3	US-08-484-223B-33	Sequence 33, App1	391	5	2.3	35	3	US-08-484-223B-107	Sequence 107, App1
319	5	2.3	35	3	US-08-484-223B-34	Sequence 34, App1	392	5	2.3	35	3	US-08-484-223B-108	Sequence 108, App1
					US-08-484-223B-35	Sequence 35, App1							

393	5	2.3	35	3	US-08-471-913A-37	Sequence 37, App1	466	5	2.3	35	4	US-08-474-349A-35	Sequence 35, App1
394	5	2.3	35	3	US-08-471-913A-38	Sequence 38, App1	467	5	2.3	35	4	US-08-474-349A-36	Sequence 36, App1
395	5	2.3	35	3	US-08-471-913A-39	Sequence 39, App1	468	5	2.3	35	4	US-08-474-349A-37	Sequence 37, App1
396	5	2.3	35	3	US-08-471-913A-40	Sequence 40, App1	469	5	2.3	35	4	US-08-474-349A-38	Sequence 38, App1
397	5	2.3	35	3	US-08-471-913A-41	Sequence 41, App1	470	5	2.3	35	4	US-08-474-349A-39	Sequence 39, App1
398	5	2.3	35	3	US-08-471-913A-42	Sequence 42, App1	471	5	2.3	35	4	US-08-474-349A-40	Sequence 40, App1
399	5	2.3	35	3	US-08-471-913A-43	Sequence 43, App1	472	5	2.3	35	4	US-08-474-349A-41	Sequence 41, App1
400	5	2.3	35	3	US-08-471-913A-44	Sequence 44, App1	473	5	2.3	35	4	US-08-474-349A-42	Sequence 42, App1
401	5	2.3	35	3	US-08-471-913A-45	Sequence 45, App1	474	5	2.3	35	4	US-08-474-349A-43	Sequence 43, App1
402	5	2.3	35	3	US-08-471-913A-46	Sequence 46, App1	475	5	2.3	35	4	US-08-474-349A-44	Sequence 44, App1
403	5	2.3	35	3	US-08-471-913A-47	Sequence 47, App1	476	5	2.3	35	4	US-08-474-349A-45	Sequence 45, App1
404	5	2.3	35	3	US-08-471-913A-48	Sequence 48, App1	477	5	2.3	35	4	US-08-474-349A-46	Sequence 46, App1
405	5	2.3	35	3	US-08-471-913A-153	Sequence 153, App1	478	5	2.3	35	4	US-08-474-349A-47	Sequence 47, App1
406	5	2.3	35	4	US-08-831-028-20	Sequence 20, App1	479	5	2.3	35	4	US-08-474-349A-153	Sequence 153, App1
407	5	2.3	35	4	US-08-974-549A-27	Sequence 27, App1	480	5	2.3	35	4	US-08-474-349A-321	Sequence 321, App1
408	5	2.3	35	4	US-08-974-549A-33	Sequence 33, App1	481	5	2.3	35	4	US-08-474-349A-322	Sequence 322, App1
409	5	2.3	35	4	US-08-905-223-331	Sequence 331, App1	482	5	2.3	35	4	US-08-474-349A-323	Sequence 323, App1
410	5	2.3	35	4	US-08-485-264A-18	Sequence 18, App1	483	5	2.3	35	4	US-08-474-349A-324	Sequence 324, App1
411	5	2.3	35	4	US-08-485-264A-33	Sequence 33, App1	484	5	2.3	35	4	US-08-474-349A-325	Sequence 325, App1
412	5	2.3	35	4	US-08-485-264A-34	Sequence 34, App1	485	5	2.3	35	4	US-08-474-349A-326	Sequence 326, App1
413	5	2.3	35	4	US-08-485-264A-35	Sequence 35, App1	486	5	2.3	35	4	US-08-474-349A-327	Sequence 327, App1
414	5	2.3	35	4	US-08-485-264A-36	Sequence 36, App1	487	5	2.3	35	4	US-08-474-349A-328	Sequence 328, App1
415	5	2.3	35	4	US-08-485-264A-37	Sequence 37, App1	488	5	2.3	35	4	US-08-474-349A-329	Sequence 329, App1
416	5	2.3	35	4	US-08-485-264A-38	Sequence 38, App1	489	5	2.3	35	4	US-08-474-349A-330	Sequence 330, App1
417	5	2.3	35	4	US-08-485-264A-39	Sequence 39, App1	490	5	2.3	35	4	US-08-474-349A-331	Sequence 331, App1
418	5	2.3	35	4	US-08-485-264A-40	Sequence 40, App1	491	5	2.3	35	4	US-08-474-349A-332	Sequence 332, App1
419	5	2.3	35	4	US-08-485-264A-41	Sequence 41, App1	492	5	2.3	35	4	US-08-474-349A-333	Sequence 333, App1
420	5	2.3	35	4	US-08-485-264A-42	Sequence 42, App1	493	5	2.3	35	4	US-08-474-349A-334	Sequence 334, App1
421	5	2.3	35	4	US-08-485-264A-43	Sequence 43, App1	494	5	2.3	35	4	US-08-474-349A-335	Sequence 335, App1
422	5	2.3	35	4	US-08-485-264A-44	Sequence 44, App1	495	5	2.3	35	4	US-08-474-349A-336	Sequence 336, App1
423	5	2.3	35	4	US-08-485-264A-45	Sequence 45, App1	496	5	2.3	35	4	US-08-474-349A-337	Sequence 337, App1
424	5	2.3	35	4	US-08-485-264A-46	Sequence 46, App1	497	5	2.3	35	4	US-08-474-349A-338	Sequence 338, App1
425	5	2.3	35	4	US-08-485-264A-47	Sequence 47, App1	498	5	2.3	35	4	US-08-474-349A-339	Sequence 339, App1
426	5	2.3	35	4	US-08-485-264A-48	Sequence 48, App1	499	5	2.3	35	4	US-08-474-349A-340	Sequence 340, App1
427	5	2.3	35	4	US-08-485-264A-153	Sequence 153, App1	500	5	2.3	35	4	US-08-474-349A-341	Sequence 341, App1
428	5	2.3	35	4	US-09-082-279B-186	Sequence 186, App1	501	5	2.3	35	4	US-08-474-349A-342	Sequence 342, App1
429	5	2.3	35	4	US-09-082-279B-79	Sequence 79, App1	502	5	2.3	35	4	US-09-315-304B-39	Sequence 39, App1
430	5	2.3	35	4	US-09-082-279B-80	Sequence 80, App1	503	5	2.3	35	4	US-09-315-304B-79	Sequence 79, App1
431	5	2.3	35	4	US-09-082-279B-187	Sequence 187, App1	504	5	2.3	35	4	US-09-315-304B-80	Sequence 80, App1
432	5	2.3	35	4	US-09-082-279B-188	Sequence 188, App1	505	5	2.3	35	4	US-09-315-304B-187	Sequence 187, App1
433	5	2.3	35	4	US-09-082-279B-189	Sequence 189, App1	506	5	2.3	35	4	US-09-315-304B-188	Sequence 188, App1
434	5	2.3	35	4	US-09-082-279B-190	Sequence 190, App1	507	5	2.3	35	4	US-09-315-304B-189	Sequence 189, App1
435	5	2.3	35	4	US-09-082-279B-191	Sequence 191, App1	508	5	2.3	35	4	US-09-315-304B-190	Sequence 190, App1
436	5	2.3	35	4	US-09-082-279B-192	Sequence 192, App1	509	5	2.3	35	4	US-09-315-304B-191	Sequence 191, App1
437	5	2.3	35	4	US-09-082-279B-193	Sequence 193, App1	510	5	2.3	35	4	US-09-315-304B-192	Sequence 192, App1
438	5	2.3	35	4	US-09-082-279B-194	Sequence 194, App1	511	5	2.3	35	4	US-09-315-304B-193	Sequence 193, App1
439	5	2.3	35	4	US-09-082-279B-195	Sequence 195, App1	512	5	2.3	35	4	US-09-315-304B-194	Sequence 194, App1
440	5	2.3	35	4	US-09-082-279B-196	Sequence 196, App1	513	5	2.3	35	4	US-09-315-304B-195	Sequence 195, App1
441	5	2.3	35	4	US-09-082-279B-197	Sequence 197, App1	514	5	2.3	35	4	US-09-315-304B-196	Sequence 196, App1
442	5	2.3	35	4	US-09-082-279B-198	Sequence 198, App1	515	5	2.3	35	4	US-09-315-304B-197	Sequence 197, App1
443	5	2.3	35	4	US-09-082-279B-199	Sequence 199, App1	516	5	2.3	35	4	US-09-315-304B-198	Sequence 198, App1
444	5	2.3	35	4	US-09-082-279B-200	Sequence 200, App1	517	5	2.3	35	4	US-09-315-304B-199	Sequence 199, App1
445	5	2.3	35	4	US-09-082-279B-201	Sequence 201, App1	518	5	2.3	35	4	US-09-315-304B-200	Sequence 200, App1
446	5	2.3	35	4	US-09-082-279B-202	Sequence 202, App1	519	5	2.3	35	4	US-09-315-304B-201	Sequence 201, App1
447	5	2.3	35	4	US-09-082-279B-203	Sequence 203, App1	520	5	2.3	35	4	US-09-315-304B-202	Sequence 202, App1
448	5	2.3	35	4	US-09-082-279B-204	Sequence 204, App1	521	5	2.3	35	4	US-09-315-304B-203	Sequence 203, App1
449	5	2.3	35	4	US-09-082-279B-205	Sequence 205, App1	522	5	2.3	35	4	US-09-315-304B-204	Sequence 204, App1
450	5	2.3	35	4	US-09-082-279B-206	Sequence 206, App1	523	5	2.3	35	4	US-09-315-304B-205	Sequence 205, App1
451	5	2.3	35	4	US-09-082-279B-207	Sequence 207, App1	524	5	2.3	35	4	US-09-315-304B-206	Sequence 206, App1
452	5	2.3	35	4	US-09-082-279B-208	Sequence 208, App1	525	5	2.3	35	4	US-09-315-304B-207	Sequence 207, App1
453	5	2.3	35	4	US-09-082-279B-209	Sequence 209, App1	526	5	2.3	35	4	US-09-315-304B-208	Sequence 208, App1
454	5	2.3	35	4	US-09-082-279B-210	Sequence 210, App1	527	5	2.3	35	4	US-09-315-304B-209	Sequence 209, App1
455	5	2.3	35	4	US-09-082-279B-211	Sequence 211, App1	528	5	2.3	35	4	US-09-315-304B-210	Sequence 210, App1
456	5	2.3	35	4	US-09-082-279B-212	Sequence 212, App1	529	5	2.3	35	4	US-09-315-304B-211	Sequence 211, App1
457	5	2.3	35	4	US-09-082-279B-213	Sequence 213, App1	530	5	2.3	35	4	US-09-315-304B-212	Sequence 212, App1
458	5	2.3	35	4	US-09-082-279B-214	Sequence 214, App1	531	5	2.3	35	4	US-09-315-304B-213	Sequence 213, App1
459	5	2.3	35	4	US-09-082-279B-215	Sequence 215, App1	532	5	2.3	35	4	US-09-315-304B-214	Sequence 214, App1
460	5	2.3	35	4	US-09-082-279B-216	Sequence 216, App1	533	5	2.3	35	4	US-09-315-304B-215	Sequence 215, App1
461	5	2.3	35	4	US-09-082-279B-217	Sequence 217, App1	534	5	2.3	35	4	US-09-315-304B-216	Sequence 216, App1
462	5	2.3	35	4	US-09-082-279B-218	Sequence 218, App1	535	5	2.3	35	4	US-09-315-304B-217	Sequence 217, App1
463	5	2.3	35	4	US-09-082-279B-219	Sequence 219, App1	536	5	2.3	35	4	US-09-315-304B-218	Sequence 218, App1
464	5	2.3	35	4	US-09-082-279B-220	Sequence 220, App1	537	5	2.3	35	4	US-09-315-304B-219	Sequence 219, App1
465	5	2.3	35	4	US-08-474-349A-34	Sequence 34, App1	538	5	2.3	35	4	US-08-722-015A-255	Sequence 255, App1

539	5	2.3	36	1	US-08-487-890A-16	Sequence 16, Appl	612	5	2.3	88	4	US-09-216-295-17	Sequence 17, Appl
540	5	2.3	36	2	US-08-478-435-16	Sequence 16, Appl	613	5	2.3	89	1	US-08-208-339A-4	Sequence 4, Appl1
541	5	2.3	36	2	US-08-337-483-16	Sequence 16, Appl	614	5	2.3	89	3	US-08-722-719-6	Sequence 6, Appl1
542	5	2.3	36	2	US-08-478-373-16	Sequence 16, Appl	615	5	2.3	90	1	US-08-464-339A-11	Sequence 11, Appl
543	5	2.3	36	3	US-08-474-671-16	Sequence 16, Appl	616	5	2.3	90	4	US-08-936-165A-262	Sequence 262, App
544	5	2.3	36	3	US-08-483-577A-16	Sequence 16, Appl	617	5	2.3	91	4	US-09-247-155-170	Sequence 170, App
545	5	2.3	36	4	US-08-897-438-16	Sequence 16, Appl	618	5	2.3	91	4	US-09-227-357-162	Sequence 162, App
546	5	2.3	36	4	US-08-637-654-16	Sequence 16, Appl	619	5	2.3	92	6	5284931-16	Patent No. 5284931
547	5	2.3	36	4	US-08-649-518-16	Sequence 16, Appl	620	5	2.3	93	4	US-08-936-165A-340	Sequence 340, App
548	5	2.3	36	4	US-09-149-476-400	Sequence 400, App	621	5	2.3	93	4	US-08-469-260A-19	Sequence 19, App
549	5	2.3	36	4	US-09-106-568E-128	Sequence 128, App	622	5	2.3	94	2	US-08-828-008-2	Sequence 2, Appl
550	5	2.3	37	1	US-08-306-871-50	Sequence 50, Appl	623	5	2.3	94	4	US-09-230-371A-21	Sequence 21, Appl
551	5	2.3	37	1	US-08-369-959-50	Sequence 50, Appl	624	5	2.3	95	4	US-09-230-637-26	Sequence 26, Appl
552	5	2.3	37	4	US-08-255-208A-18	Sequence 18, Appl	625	5	2.3	96	5	PCT-US93-08744-4	Sequence 4, Appl1
553	5	2.3	43	1	US-08-428-091-2	Sequence 2, Appl1	626	5	2.3	97	4	PCT-US93-08744-4	Sequence 500, Ap
554	5	2.3	43	4	US-08-988-856B-21	Sequence 21, Appl	627	5	2.3	98	1	US-09-134-001C-5300	Sequence 4, Appl1
555	5	2.3	43	4	US-09-445-247-16	Sequence 16, Appl	628	5	2.3	98	2	US-09-053-945-4	Sequence 4, Appl1
556	5	2.3	44	1	US-08-118-270-305	Sequence 305, App	629	5	2.3	98	3	US-09-205-680A-3	Sequence 3, Appl1
557	5	2.3	44	1	US-08-118-270-306	Sequence 306, App	630	5	2.3	98	3	US-09-048-889-11	Sequence 11, Appl
558	5	2.3	44	2	US-08-598-873-23	Sequence 23, Appl	631	5	2.3	98	3	US-08-905-223-484	Sequence 484, App
559	5	2.3	44	4	US-08-605-430-23	Sequence 23, Appl	632	5	2.3	100	4	US-08-473-981A-11	Sequence 11, Appl
560	5	2.3	44	4	US-08-988-856B-6	Sequence 6, Appl	633	5	2.3	100	2	US-08-474-087-11	Sequence 11, Appl
561	5	2.3	44	4	US-08-988-856B-16	Sequence 16, Appl	634	5	2.3	100	2	US-08-876-991-13	Sequence 13, Appl
562	5	2.3	44	5	PCT-US93-08528-305	Sequence 305, App	635	5	2.3	100	2	US-09-059-853-13	Sequence 13, Appl
563	5	2.3	44	5	PCT-US93-08528-306	Sequence 306, App	636	5	2.3	100	2	US-09-245-041-126	Sequence 126, App
564	5	2.3	45	1	US-08-118-270-307	Sequence 307, App	637	5	2.3	101	1	US-08-220-401-5	Sequence 5, Appl1
565	5	2.3	45	1	PCT-US93-08528-307	Sequence 307, App	638	5	2.3	101	2	US-08-437-362-5	Sequence 5, Appl1
566	5	2.3	46	1	US-08-415-751-10	Sequence 10, Appl	639	5	2.3	102	4	US-09-280-598-50	Sequence 50, Appl
567	5	2.3	47	1	US-08-118-270-308	Sequence 308, App	640	5	2.3	102	4	US-09-308-188-1	Sequence 11, Appl
568	5	2.3	47	4	US-09-149-476-359	Sequence 359, App	641	5	2.3	103	1	US-08-276-852-112	Sequence 112, App
569	5	2.3	47	4	US-09-149-476-364	Sequence 364, App	642	5	2.3	104	1	US-08-899-575-112	Sequence 112, App
570	5	2.3	47	5	PCT-US93-08528-308	Sequence 308, App	643	5	2.3	104	1	US-08-899-575-112	Sequence 112, App
571	5	2.3	50	4	US-09-177-249-167	Sequence 167, App	644	5	2.3	104	3	US-08-946-329A-100	Sequence 100, App
572	5	2.3	53	4	US-09-257-179-44	Sequence 44, App	645	5	2.3	104	4	US-09-146-580-2	Sequence 2, Appl1
573	5	2.3	56	1	US-07-752-101A-16	Sequence 16, Appl	646	5	2.3	104	4	US-09-486-382B-4	Sequence 4, Appl1
574	5	2.3	56	2	US-08-245-511-45	Sequence 45, Appl	647	5	2.3	104	5	PCT-US95-08743-112	Sequence 112, App
575	5	2.3	56	2	US-08-600-993A-45	Sequence 45, Appl	648	5	2.3	106	1	US-07-803-623B-7	Sequence 7, Appl1
576	5	2.3	56	3	US-08-486-099-101	Sequence 101, App	649	5	2.3	106	2	US-08-806-084-7	Sequence 7, Appl1
577	5	2.3	56	3	US-08-360-107A-90	Sequence 90, Appl	650	5	2.3	106	2	US-08-911-319A-3	Sequence 3, Appl1
578	5	2.3	56	3	US-08-360-107A-111	Sequence 111, App	651	5	2.3	106	3	US-09-352-619-7	Sequence 7, Appl1
579	5	2.3	56	3	US-08-460-223B-101	Sequence 101, App	652	5	2.3	106	4	US-09-201-870A-7	Sequence 7, Appl1
580	5	2.3	56	3	US-08-919-587-101	Sequence 101, App	653	5	2.3	106	4	US-09-162-564-2	Sequence 2, Appl1
581	5	2.3	56	3	US-08-475-668A-101	Sequence 101, App	654	5	2.3	106	4	US-09-162-564-4	Sequence 4, Appl1
582	5	2.3	56	3	US-08-485-551A-101	Sequence 101, App	655	5	2.3	107	4	US-08-654-482-7	Sequence 7, Appl1
583	5	2.3	56	3	US-08-471-913A-101	Sequence 101, App	656	5	2.3	109	1	US-08-466-886-27	Sequence 27, App
584	5	2.3	56	4	US-08-485-264A-101	Sequence 101, App	657	5	2.3	109	4	US-08-466-886-27	Sequence 27, App
585	5	2.3	56	4	US-08-474-349A-101	Sequence 101, App	658	5	2.3	112	4	US-08-899-330-7	Sequence 7, Appl1
586	5	2.3	56	4	US-08-255-208A-35	Sequence 35, Appl	659	5	2.3	113	1	US-08-241-853-15	Sequence 15, Appl
587	5	2.3	57	4	US-08-630-915A-209	Sequence 209, App	660	5	2.3	113	2	US-08-850-917-15	Sequence 15, Appl
588	5	2.3	62	1	US-07-662-005A-14	Sequence 14, Appl	661	5	2.3	113	4	US-09-311-352B-2	Sequence 2, Appl1
589	5	2.3	62	1	US-09-295-996B-4	Sequence 4, Appl1	662	5	2.3	114	4	US-09-188-930-161	Sequence 161, App
590	5	2.3	64	1	US-08-428-091-5	Sequence 5, Appl1	663	5	2.3	114	4	US-09-188-930-288	Sequence 288, App
591	5	2.3	65	1	US-08-137-800-46	Sequence 46, Appl	664	5	2.3	117	1	US-07-752-101A-37	Sequence 37, Appl
592	5	2.3	65	1	US-08-477-383-46	Sequence 46, Appl	665	5	2.3	117	1	US-07-752-101A-40	Sequence 40, Appl
593	5	2.3	65	1	US-08-487-174-46	Sequence 46, Appl	666	5	2.3	117	1	US-07-752-101A-42	Sequence 42, Appl
594	5	2.3	65	1	US-08-480-750-46	Sequence 46, Appl	667	5	2.3	117	1	US-08-249-103-6	Sequence 6, Appl1
595	5	2.3	67	6	US-08-450-360-10	Sequence 10, Appl	668	5	2.3	117	2	US-08-886-863-6	Sequence 6, Appl1
596	5	2.3	67	6	5320958-14	Patent No. 5320958	669	5	2.3	117	2	US-08-886-863-6	Sequence 6, Appl1
597	5	2.3	68	2	US-08-385-580A-20	Sequence 20, Appl	670	5	2.3	117	2	US-08-901-206-5	Sequence 5, Appl1
598	5	2.3	68	3	US-09-021-520-40	Sequence 20, Appl	671	5	2.3	117	4	US-09-175-229-6	Sequence 6, Appl1
599	5	2.3	70	4	US-09-367-953B-84	Sequence 84, Appl	672	5	2.3	117	5	PCT-US95-06764-6	Sequence 6, Appl1
600	5	2.3	73	4	US-09-149-476-579	Sequence 579, App	673	5	2.3	118	3	US-09-065-059-5	Sequence 5, Appl1
601	5	2.3	76	2	US-08-469-537A-38	Sequence 38, Appl	674	5	2.3	118	4	US-09-045-764A-17	Sequence 17, App
602	5	2.3	76	2	US-08-469-537A-39	Sequence 39, Appl	675	5	2.3	119	4	US-09-202-712A-5	Sequence 17, App
603	5	2.3	76	2	US-08-469-537A-40	Sequence 40, Appl	676	5	2.3	119	4	US-09-056-556-197	Sequence 197, App
604	5	2.3	77	4	US-08-905-223-474	Sequence 474, App	677	5	2.3	119	4	US-09-072-596-192	Sequence 192, App
605	5	2.3	78	1	US-08-375-346A-6	Sequence 6, Appl1	678	5	2.3	120	1	US-07-929-206-2	Sequence 2, Appl1
606	5	2.3	78	2	US-08-467-123B-6	Sequence 6, Appl1	679	5	2.3	120	2	US-08-313-185-43	Sequence 43, Appl
607	5	2.3	78	2	US-08-576-626A-37	Sequence 37, Appl	680	5	2.3	120	3	US-08-459-499-2	Sequence 2, Appl1
608	5	2.3	78	4	US-08-936-165A-503	Sequence 503, App	681	5	2.3	122	4	US-09-082-614A-43	Sequence 43, Appl
609	5	2.3	79	2	US-09-196-048-2	Sequence 2, Appl	682	5	2.3	123	1	US-09-237-357-285	Sequence 285, App
610	5	2.3	83	2	US-08-537-811-47	Sequence 47, Appl	683	5	2.3	123	1	US-09-053-945-2	Sequence 2, Appl1
611	5	2.3	84	2	US-08-031-538-67	Sequence 67, Appl	684	5	2.3	123	2	US-08-061-514A-2	Sequence 2, Appl1
												US-08-482-728A-6	Sequence 6, Appl1

685	2.3	124	3	US-08-865-297-4	Sequence 4, Appli	758	5	2.3	170	4	US-08-817-441-65	Sequence 65, Appli
686	2.3	124	3	US-08-904-871-8	Sequence 8, Appli	759	5	2.3	171	1	US-08-313-075A-32	Sequence 32, Appli
687	2.3	125	4	US-08-818-112-84	Sequence 84, Appli	760	5	2.3	171	1	US-08-725-459B-14	Sequence 14, Appli
688	2.3	125	4	US-08-818-111-85	Sequence 85, Appli	761	5	2.3	171	1	US-09-134-001C-3566	Sequence 3566, Ap
689	2.3	125	4	US-09-055-556-84	Sequence 84, Appli	762	5	2.3	171	4	US-08-817-441-96	Sequence 96, Appli
690	2.3	125	4	US-09-072-596-85	Sequence 85, Appli	763	5	2.3	172	4	US-09-134-001C-5349	Sequence 5349, Ap
691	2.3	127	4	US-09-146-580-1	Sequence 1, Appli	764	5	2.3	172	4	US-09-704-139-5	Sequence 5, Appli
692	2.3	128	1	US-07-855-411A-1	Sequence 1, Appli	765	5	2.3	173	4	US-09-357-251-88	Sequence 2, Appli
693	2.3	128	1	US-07-855-411A-2	Sequence 2, Appli	766	5	2.3	173	4	US-08-936-165A-476	Sequence 476, App
694	2.3	130	1	US-08-455-550-10	Sequence 10, Appli	767	5	2.3	173	4	US-09-704-139-4	Sequence 4, Appli
695	2.3	130	4	US-09-485-316A-11	Sequence 11, Appli	768	5	2.3	174	4	US-09-149-476-424	Sequence 424, App
696	2.3	130	4	US-09-485-316A-12	Sequence 12, Appli	769	5	2.3	175	4	US-08-973-961-7	Sequence 7, Appli
697	2.3	130	4	US-09-485-316A-13	Sequence 13, Appli	770	5	2.3	175	4	US-08-973-961-8	Sequence 8, Appli
698	2.3	131	3	US-09-102-977-2	Sequence 2, Appli	771	5	2.3	176	1	US-09-376-113-7	Sequence 7, Appli
699	2.3	131	4	US-09-034-088A-2	Sequence 2, Appli	772	5	2.3	176	1	US-08-414-926A-18	Sequence 18, Appli
700	2.3	132	2	US-08-933-750C-38	Sequence 38, Appli	773	5	2.3	176	2	US-08-926-922-18	Sequence 18, Appli
701	2.3	132	4	US-09-234-613-38	Sequence 38, Appli	774	5	2.3	176	3	US-09-253-682-18	Sequence 18, Appli
702	2.3	132	4	US-09-034-088A-4	Sequence 4, Appli	775	5	2.3	176	3	US-09-527-657-18	Sequence 18, Appli
703	2.3	135	1	US-08-336-618-20	Sequence 20, Appli	776	5	2.3	177	2	US-08-850-910A-18	Sequence 18, Appli
704	2.3	136	1	US-09-171-456-16	Sequence 16, Appli	777	5	2.3	177	4	US-09-820-920-3	Sequence 20, App
705	2.3	136	4	US-09-589-287B-20	Sequence 20, Appli	778	5	2.3	177	4	US-09-370-838-402	Sequence 208, App
706	2.3	137	3	US-08-875-540-12	Sequence 12, Appli	779	5	2.3	177	4	US-08-934-959-5	Sequence 5, Appli
707	2.3	137	4	US-09-473-634-12	Sequence 12, Appli	780	5	2.3	178	2	US-08-122-443-9	Sequence 9, Appli
708	2.3	137	4	US-09-201-227A-6	Sequence 6, Appli	781	5	2.3	178	3	US-08-765-094C-24	Sequence 24, Appli
709	2.3	137	4	US-09-201-227A-19	Sequence 19, Appli	782	5	2.3	178	4	US-09-082-797-24	Sequence 24, Appli
710	2.3	138	4	US-09-615-192A-19	Sequence 384, App	783	5	2.3	178	4	US-09-315-689-5	Sequence 5, Appli
711	2.3	139	1	US-08-542-363-31	Sequence 31, Appli	784	5	2.3	178	4	US-09-134-001C-4767	Sequence 4767, Ap
712	2.3	139	4	US-09-100-089-31	Sequence 31, Appli	785	5	2.3	179	4	US-09-178-973B-15	Sequence 15, Appli
713	2.3	139	4	US-09-134-001C-5327	Sequence 31, Appli	786	5	2.3	179	4	US-09-178-973B-16	Sequence 16, Appli
714	2.3	139	4	US-09-670-827-31	Sequence 31, Appli	787	5	2.3	179	4	US-09-309-317-4	Sequence 4, Appli
715	2.3	140	1	US-07-285-515-2	Sequence 2, Appli	788	5	2.3	179	4	US-09-419-568F-27	Sequence 27, Appli
716	2.3	142	1	US-07-661-610C-3	Sequence 3, Appli	789	5	2.3	179	4	US-09-419-568F-28	Sequence 28, Appli
717	2.3	142	2	US-08-448-438-1	Sequence 2, Appli	790	5	2.3	179	4	US-09-354-243B-27	Sequence 27, Appli
718	2.3	142	2	US-08-448-438-2	Sequence 2, Appli	791	5	2.3	179	4	US-09-354-243B-28	Sequence 28, Appli
719	2.3	142	2	US-08-448-438-3	Sequence 3, Appli	792	5	2.3	179	4	US-09-615-182A-376	Sequence 376, App
720	2.3	142	2	US-08-448-438-4	Sequence 4, Appli	793	5	2.3	179	4	US-08-874-832-15	Sequence 15, Appli
721	2.3	143	3	US-08-946-329A-52	Sequence 52, Appli	794	5	2.3	180	2	US-09-097-233-15	Sequence 15, Appli
722	2.3	143	3	US-09-129-075-12	Sequence 12, Appli	795	5	2.3	180	3	US-08-936-165A-337	Sequence 337, App
723	2.3	143	4	US-09-544-123-12	Sequence 12, Appli	796	5	2.3	180	4	US-09-134-001C-3947	Sequence 3947, Ap
724	2.3	145	2	US-08-788-58A-1	Sequence 1, Appli	797	5	2.3	180	4	US-09-615-192A-281	Sequence 281, App
725	2.3	145	2	US-08-708-541A-28	Sequence 28, Appli	798	5	2.3	180	4	US-09-465-558-2	Sequence 2, Appli
726	2.3	145	2	US-08-708-541A-32	Sequence 32, Appli	799	5	2.3	182	1	US-08-626-264-28	Sequence 28, Appli
727	2.3	147	4	US-08-883-086-3	Sequence 3, Appli	800	5	2.3	182	2	US-08-698-805-4	Sequence 4, Appli
728	2.3	149	2	US-08-606-143-24	Sequence 24, Appli	801	5	2.3	182	2	US-09-561-500-14	Sequence 14, Appli
729	2.3	149	3	US-08-338-579A-95	Sequence 95, Appli	802	5	2.3	182	4	US-09-561-108-14	Sequence 14, Appli
730	2.3	152	3	US-08-916-043-2	Sequence 2, Appli	803	5	2.3	182	4	US-09-315-689-3	Sequence 3, Appli
731	2.3	152	3	US-09-357-251-30	Sequence 30, Appli	804	5	2.3	182	4	US-09-561-526-14	Sequence 14, Appli
732	2.3	154	6	5212286-6	Patent No. 5212286	805	5	2.3	182	4	US-09-206-059-2	Sequence 8, Appli
733	2.3	154	3	US-08-630-172-7	Sequence 7, Appli	806	5	2.3	183	4	US-09-105-343A-8	Sequence 13, Appli
734	2.3	154	4	US-09-375-419-7	Sequence 3873, App	807	5	2.3	183	4	US-09-475-316A-13	Sequence 4105, Ap
735	2.3	154	4	US-09-134-001C-3873	Sequence 210, App	808	5	2.3	185	6	5489533-4	Sequence 2998, App
736	2.3	156	4	US-09-370-838-210	Sequence 6, Appli	809	5	2.3	186	4	US-09-134-001C-4998	Sequence 21, Appli
737	2.3	156	4	US-09-173-300-6	Sequence 3371, Ap	810	5	2.3	187	4	US-08-464-517-21	Sequence 21, Appli
738	2.3	157	4	US-07-906-871-10	Sequence 10, Appli	811	5	2.3	189	2	US-08-246-361A-21	Sequence 2, Appli
739	2.3	158	1	US-07-906-871-12	Sequence 12, Appli	812	5	2.3	189	3	US-08-865-297-21	Sequence 21, Appli
740	2.3	159	4	US-09-149-475-427	Sequence 427, App	813	5	2.3	189	3	US-08-463-772-21	Sequence 21, Appli
741	2.3	160	2	US-08-448-438-5	Sequence 5, Appli	814	5	2.3	189	3	US-09-384-162-14	Sequence 14, Appli
742	2.3	160	2	US-08-448-438-6	Sequence 6, Appli	815	5	2.3	189	4	PCT-US93-05000-21	Sequence 25, Appli
743	2.3	160	2	US-08-448-438-7	Sequence 7, Appli	816	5	2.3	189	5	US-07-885-689A-29	Sequence 29, Appli
744	2.3	160	2	US-08-448-438-8	Sequence 8, Appli	817	5	2.3	190	1	US-08-093-383-3	Sequence 3, Appli
745	2.3	161	4	US-08-955-629C-2	Sequence 2, Appli	818	5	2.3	191	1	US-08-286-819A-21	Sequence 21, Appli
746	2.3	161	4	US-08-858-207A-520	Sequence 520, App	819	5	2.3	191	2	US-08-980-357-21	Sequence 21, Appli
747	2.3	166	4	US-09-452-624A-2	Sequence 2, Appli	820	5	2.3	191	4	US-08-858-207A-774	Sequence 339, App
748	2.3	166	4	US-09-006-428A-3	Sequence 3, Appli	821	5	2.3	191	4	US-08-383-621-2	Sequence 2, Appli
749	2.3	168	3	US-08-669-408A-4	Sequence 25, Appli	822	5	2.3	191	4	US-08-459-906-2	Sequence 4, Appli
750	2.3	168	4	US-08-580-031A-25	Sequence 5, Appli	823	5	2.3	193	2	US-08-233-005-4	Sequence 2, Appli
751	2.3	168	4	US-09-376-113-5	Sequence 4, Appli	824	5	2.3	194	1	US-08-428-943-4	Sequence 4, Appli
752	2.3	169	2	US-08-895-939-4	Sequence 48, Appli	825	5	2.3	194	1	US-09-016-644-8	Sequence 4, Appli
753	2.3	169	4	US-08-444-818-48	Sequence 2948, Ap	826	5	2.3	194	5	PCT-US95-04858-4	Sequence 4, Appli
754	2.3	169	4	US-09-188-820-4	Sequence 31, Appli	827	5	2.3	830			
755	2.3	169	4	US-09-134-001C-2948		828	5					
756	2.3	170	3	US-08-725-459B-31		829	5					
757	2.3	170	3	US-08-725-459B-31		830	5					



831	5	2.3	197	4	US-09-068-195-17	Sequence 17, Appl	504	5	2.3	218	4	US-08-256-847C-3	Sequence 3, Appl
832	5	2.3	198	4	US-09-134-001C-4395	Sequence 4395, Ap	505	5	2.3	219	1	US-08-961-083-28	Sequence 28, Appl
833	5	2.3	199	1	US-07-801-164A-9	Sequence 4, Appl	906	5	2.3	221	1	US-08-698-978-2	Sequence 2, Appl
834	5	2.3	199	2	US-08-665-647-9	Sequence 9, Appl	907	5	2.3	221	4	US-08-470-953A-3	Sequence 3, Appl
835	5	2.3	199	4	US-09-213-293D-7	Sequence 7, Appl	908	5	2.3	222	1	US-07-668-952C-46	Sequence 46, Appl
836	5	2.3	200	2	US-08-531-525-33	Sequence 33, Appl	909	5	2.3	222	2	US-08-467-948A-27	Sequence 27, Appl
837	5	2.3	200	2	US-08-718-270A-33	Sequence 33, Appl	910	5	2.3	222	3	US-08-467-947A-27	Sequence 27, Appl
838	5	2.3	200	4	US-09-068-740A-2	Sequence 2, Appl	911	5	2.3	222	5	PCT-US91-02714-43	Sequence 43, Appl
839	5	2.3	200	6	5189147-5	Patent No. 5189147	912	5	2.3	222	5	PCT-US91-02714-43	Sequence 43, Appl
840	5	2.3	201	3	US-08-725-459B-33	Sequence 33, Appl	913	5	2.3	223	2	US-08-484-993B-49	Sequence 49, Appl
841	5	2.3	201	4	US-09-411-339C-1	Sequence 1, Appl	914	5	2.3	223	2	US-08-484-158B-49	Sequence 49, Appl
842	5	2.3	201	4	US-09-411-335-1	Sequence 1, Appl	915	5	2.3	223	2	US-08-484-596A-49	Sequence 49, Appl
843	5	2.3	201	4	US-09-466-276-1	Sequence 1, Appl	916	5	2.3	223	2	US-08-480-150A-49	Sequence 49, Appl
844	5	2.3	202	1	US-07-803-623B-8	Sequence 8, Appl	917	5	2.3	223	3	US-08-458-331-49	Sequence 49, Appl
845	5	2.3	202	1	US-08-531-525-24	Sequence 24, Appl	918	5	2.3	223	3	US-08-149-223A-49	Sequence 49, Appl
846	5	2.3	202	2	US-08-806-084-8	Sequence 8, Appl	919	5	2.3	223	4	US-09-124-338A-34	Sequence 34, Appl
847	5	2.3	202	2	US-08-718-270A-24	Sequence 24, Appl	920	5	2.3	223	4	US-09-771-975-34	Sequence 34, Appl
848	5	2.3	202	4	US-09-411-329C-16	Sequence 16, Appl	921	5	2.3	224	1	US-08-287-959-5	Sequence 5, Appl
849	5	2.3	202	4	US-09-201-970A-8	Sequence 8, Appl	922	5	2.3	224	2	US-08-924-759-22	Sequence 22, Appl
850	5	2.3	203	2	US-08-531-525-23	Sequence 23, Appl	923	5	2.3	224	3	US-09-248-335-22	Sequence 22, Appl
851	5	2.3	203	2	US-08-718-270A-23	Sequence 23, Appl	924	5	2.3	224	4	US-08-944-483-35	Sequence 35, Appl
852	5	2.3	203	4	US-09-124-141-9	Sequence 9, Appl	925	5	2.3	226	4	US-09-134-001C-5562	Sequence 5562, Ap
853	5	2.3	203	4	US-09-411-329C-5	Sequence 5, Appl	926	5	2.3	228	4	US-09-134-001C-4694	Sequence 4694, Ap
854	5	2.3	203	4	US-09-411-329C-15	Sequence 15, Appl	927	5	2.3	228	4	US-09-134-001C-5495	Sequence 5495, Ap
855	5	2.3	203	4	US-09-411-335-3	Sequence 3, Appl	928	5	2.3	228	4	US-09-182-145-19	Sequence 19, Appl
856	5	2.3	203	4	US-09-466-276-3	Sequence 3, Appl	929	5	2.3	229	4	US-08-936-165A-519	Sequence 519, Appl
857	5	2.3	204	2	US-08-208-005C-5	Sequence 5, Appl	930	5	2.3	229	4	US-09-182-145-99	Sequence 99, Appl
858	5	2.3	204	2	US-08-531-525-32	Sequence 32, Appl	931	5	2.3	230	4	US-09-182-145-98	Sequence 98, Appl
859	5	2.3	204	2	US-08-808-550-32	Sequence 5, Appl	932	5	2.3	231	2	US-08-622-354-5	Sequence 5, Appl
860	5	2.3	204	2	US-09-038-597A-5	Sequence 5, Appl	933	5	2.3	231	3	US-08-725-459B-32	Sequence 32, Appl
861	5	2.3	204	2	US-08-718-270A-32	Sequence 32, Appl	934	5	2.3	231	3	US-08-721-259-2	Sequence 2, Appl
862	5	2.3	204	2	US-08-576-626A-40	Sequence 40, Appl	935	5	2.3	231	4	US-08-974-380-2	Sequence 2, Appl
863	5	2.3	205	1	US-07-870-029-2	Sequence 2, Appl	936	5	2.3	231	4	US-09-182-145-97	Sequence 97, Appl
864	5	2.3	205	1	US-08-233-005-2	Sequence 2, Appl	937	5	2.3	231	4	US-09-621-625A-30	Sequence 30, Appl
865	5	2.3	205	1	US-08-428-943-2	Sequence 2, Appl	938	5	2.3	231	4	US-09-621-625A-32	Sequence 32, Appl
866	5	2.3	205	2	US-08-531-525-31	Sequence 31, Appl	939	5	2.3	231	4	US-09-530-9423-2	Sequence 23, Appl
867	5	2.3	205	2	US-08-729-152-8	Sequence 8, Appl	940	5	2.3	232	2	US-08-446-345-38	Sequence 38, Appl
868	5	2.3	205	2	US-08-718-270A-31	Sequence 31, Appl	941	5	2.3	232	4	US-09-182-145-96	Sequence 96, Appl
869	5	2.3	205	4	US-09-016-649-2	Sequence 2, Appl	942	5	2.3	233	1	US-08-044-621D-24	Sequence 24, Appl
870	5	2.3	205	4	US-08-818-112-80	Sequence 80, Appl	943	5	2.3	233	4	US-08-818-112-76	Sequence 76, Appl
871	5	2.3	205	4	US-09-286-529-5	Sequence 5, Appl	944	5	2.3	233	4	US-08-818-111-77	Sequence 77, Appl
872	5	2.3	205	4	US-08-818-111-81	Sequence 81, Appl	945	5	2.3	233	4	US-09-056-556-76	Sequence 76, Appl
873	5	2.3	205	4	US-09-056-556-80	Sequence 80, Appl	946	5	2.3	233	4	US-09-182-145-95	Sequence 95, Appl
874	5	2.3	205	4	US-09-072-596-81	Sequence 81, Appl	947	5	2.3	233	4	US-09-072-596-77	Sequence 77, Appl
875	5	2.3	205	5	PCT-US95-04858-2	Sequence 2, Appl	948	5	2.3	234	4	US-09-182-145-94	Sequence 94, Appl
876	5	2.3	206	3	US-09-078-317-15	Sequence 15, Appl	949	5	2.3	234	4	US-09-157-664-2	Sequence 2, Appl
877	5	2.3	206	4	US-09-454-818-15	Sequence 15, Appl	950	5	2.3	235	3	US-08-807-151-3	Sequence 3, Appl
878	5	2.3	208	1	US-08-309-512-4	Sequence 4, Appl	951	5	2.3	235	4	US-09-478-957-3	Sequence 3, Appl
879	5	2.3	208	1	US-08-680-726A-72	Sequence 72, Appl	952	5	2.3	236	2	US-09-182-145-93	Sequence 93, Appl
880	5	2.3	208	2	US-08-531-525-15	Sequence 15, Appl	953	5	2.3	236	2	US-08-464-617-22	Sequence 22, Appl
881	5	2.3	208	2	US-08-531-525-17	Sequence 17, Appl	954	5	2.3	236	2	US-08-464-616A-22	Sequence 22, Appl
882	5	2.3	208	2	US-08-718-270A-15	Sequence 15, Appl	955	5	2.3	236	3	US-08-657-713-1	Sequence 1, Appl
883	5	2.3	208	2	US-08-718-270A-17	Sequence 17, Appl	956	5	2.3	236	3	US-08-463-772-22	Sequence 22, Appl
884	5	2.3	208	4	US-09-092-409-72	Sequence 72, Appl	957	5	2.3	236	3	US-08-493-071-15	Sequence 15, Appl
885	5	2.3	208	4	US-08-097-869-6	Sequence 6, Appl	958	5	2.3	236	4	US-09-182-145-92	Sequence 92, Appl
886	5	2.3	208	5	PCT-US92-08156A-4	Sequence 4, Appl	959	5	2.3	236	5	PCT-US93-05000-22	Sequence 22, Appl
887	5	2.3	210	2	US-08-531-525-16	Sequence 16, Appl	960	5	2.3	237	2	US-08-464-617-24	Sequence 24, Appl
888	5	2.3	210	2	US-08-718-270A-16	Sequence 16, Appl	961	5	2.3	237	2	US-08-246-616A-24	Sequence 24, Appl
889	5	2.3	210	4	US-08-961-083-36	Sequence 36, Appl	962	5	2.3	237	3	US-09-216-001-1	Sequence 1, Appl
890	5	2.3	210	4	US-09-247-155-117	Sequence 117, App	963	5	2.3	237	3	US-08-388-553-642	Sequence 642, App
891	5	2.3	211	1	US-08-709-912-6	Sequence 6, Appl	964	5	2.3	237	3	US-08-488-551B-642	Sequence 642, App
892	5	2.3	211	2	US-09-047-370-6	Sequence 6, Appl	965	5	2.3	237	3	US-08-463-772-24	Sequence 24, Appl
893	5	2.3	211	4	US-09-130-491-4	Sequence 4, Appl	966	5	2.3	237	4	US-08-878-862-1	Sequence 1, Appl
894	5	2.3	212	2	US-08-531-525-18	Sequence 18, Appl	967	5	2.3	237	4	US-09-134-001C-5561	Sequence 5531, Ap
895	5	2.3	212	2	US-08-718-270A-18	Sequence 18, Appl	968	5	2.3	237	4	US-09-182-145-91	Sequence 91, Appl
896	5	2.3	212	4	US-09-399-913-67	Sequence 67, Appl	969	5	2.3	237	5	PCT-US93-05000-24	Sequence 24, Appl
897	5	2.3	213	3	US-08-718-738-18	Sequence 18, Appl	970	5	2.3	238	4	US-09-182-145-90	Sequence 90, Appl
898	5	2.3	213	4	US-09-068-195-16	Sequence 16, Appl	971	5	2.3	239	4	US-08-913-014A-2	Sequence 2, Appl
899	5	2.3	213	4	US-09-221-844-18	Sequence 18, Appl	972	5	2.3	239	4	US-08-913-014A-3	Sequence 3, Appl
900	5	2.3	214	4	US-09-411-578-1	Sequence 1, Appl	973	5	2.3	239	4	US-09-182-145-89	Sequence 89, Appl
901	5	2.3	216	4	US-08-952-127-22	Sequence 22, Appl	974	5	2.3	240	3	US-08-942-001-9	Sequence 9, Appl
902	5	2.3	216	4	US-09-465-558-46	Sequence 46, Appl	975	5	2.3	240	4	US-09-337-386-9	Sequence 9, Appl
903	5	2.3	217	2	US-09-105-651-2	Sequence 2, Appl	976	5	2.3	240	4	US-09-846-922-9	Sequence 9, Appl



977 5 2.3 240 4 US-09-182-145-88 Sequence 88, Appl  
978 5 2.3 241 4 US-09-182-145-87 Sequence 87, Appl  
979 5 2.3 242 4 US-09-182-145-86 Sequence 86, Appl  
980 5 2.3 243 4 US-08-858-207A-338 Sequence 338, App  
981 5 2.3 243 4 US-09-182-145-85 Sequence 85, Appl  
982 5 2.3 244 1 US-08-318-947A-21 Sequence 21, Appl  
983 5 2.3 244 2 US-08-463-911-7 Sequence 21, Appl  
984 5 2.3 244 2 US-08-795-303-21 Sequence 21, Appl  
985 5 2.3 244 4 US-09-140-804-3 Sequence 84, Appl  
986 5 2.3 244 4 US-09-182-145-84 Sequence 20, Appl  
987 5 2.3 244 4 US-09-336-535-20 Sequence 1, Appl  
988 5 2.3 244 4 US-09-530-423-1 Sequence 3, Appl  
989 5 2.3 245 4 US-09-163-507-3 Sequence 83, Appl  
990 5 2.3 245 4 US-09-182-145-83 Sequence 82, Appl  
991 5 2.3 246 4 US-09-182-145-82 Sequence 2, Appl  
992 5 2.3 247 2 US-08-463-911-2 Sequence 2, Appl  
993 5 2.3 247 2 US-08-956-267A-2 Sequence 19, Appl  
994 5 2.3 247 4 US-08-501-126-19 Sequence 8, Appl  
995 5 2.3 247 4 US-09-140-804-8 Sequence 7, Appl  
996 5 2.3 247 4 US-09-118-408-3 Sequence 16, Appl  
997 5 2.3 247 4 US-09-230-196-16 Sequence 81, Appl  
998 5 2.3 247 4 US-09-182-145-81 Sequence 4, Appl  
999 5 2.3 247 4 US-09-157-864-4 Sequence 3, Appl  
1000 5 2.3 247 4 US-09-506-855-3

## ALIGNMENTS

RESULT 1  
US-09-196-293-15  
; Sequence 15, Application US/09196293  
; Patent No. 6183755  
; GENERAL INFORMATION:  
; APPLICANT: Fuchs, Renate  
; APPLICANT: Motz, Manfred  
; APPLICANT: Soutcheck, Erwin  
; APPLICANT: Wilske, Bettina  
; APPLICANT: Preac-Mursic, Vera  
; TITLE OF INVENTION: Active proteins from Borrelia  
; FILE REFERENCE: burgdorferi  
; CURRENT APPLICATION NUMBER: US/09/196,293  
; EARLIER FILING DATE: 1998-11-19  
; EARLIER APPLICATION NUMBER: US 08/209,603  
; EARLIER FILING DATE: 1994-03-10  
; EARLIER APPLICATION NUMBER: US 07/862,535  
; EARLIER FILING DATE: 1992-06-19  
; EARLIER APPLICATION NUMBER: WO PCT/EP90/02282  
; EARLIER FILING DATE: 1990-12-21  
; EARLIER APPLICATION NUMBER: DE P39 42 728.5  
; EARLIER FILING DATE: 1989-12-22  
; EARLIER APPLICATION NUMBER: DE P40 18 988.0  
; EARLIER FILING DATE: 1990-06-13  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15  
; LENGTH: 209  
; TYPE: PRT  
; ORGANISM: Borrelia burgdorferi  
; US-09-196-293-15

Query Match 3.2%; Score 7; DB 4; Length 209;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 AKEALIK 45  
|||||  
DB 155 AKEALIK 161

RESULT 2  
US-08-235-836C-32

; Sequence 32, Application US/08235836C  
; Patent No. 6248562  
; GENERAL INFORMATION:  
; APPLICANT: Dunn, John J.  
; APPLICANT: Luft, Benjamin J.  
; TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising  
; NUMBER OF INVENTION: Borrelia Polypeptides and Uses Therefor  
; CORRESPONDENCE ADDRESS: 144  
; ADDRESS: Brookhaven National Laboratory  
; STREET:  
; CITY: Upton  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 11973  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/235,836C  
; FILING DATE: 29-APR-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/148,191  
; FILING DATE: 01-11-93  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bogosian, Margaret C.  
; REGISTRATION NUMBER: 25,324  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 282-7338  
; TELEFAX: (516) 282-3729  
; INFORMATION FOR SEQ ID NO: 32:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 209 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-235-836C-32

Query Match 3.2%; Score 7; DB 4; Length 209;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 AKEALIK 45  
|||||  
DB 154 AKEALIK 160

RESULT 3  
US-08-158-353-3  
; Sequence 3, Application US/08158353  
; Patent No. 5620862  
; GENERAL INFORMATION:  
; APPLICANT: Padula, Steven J.  
; TITLE OF INVENTION: Methods for Diagnosing Early Lyme  
; TITLE OF INVENTION: Disease  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.  
; STREET: Two Militia Drive  
; CITY: Lexington  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02173  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/158.353  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Carroll, Alice O.  
REGISTRATION NUMBER: 33,542  
REFERENCE/DOCKET NUMBER: UCT93-05  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-9540  
TELEFAX: 617-861-6240  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 210 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-158-353-3

Query Match 3.2%; Score 7; DB 1; Length 210;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 AKEALIK 45  
Db 155 AKEALIK 161

RESULT 4  
US-08-209-603E-15  
Sequence 15, Application US/08209603E  
Patent No. 6248538  
GENERAL INFORMATION:  
APPLICANT: FUCHS, RENATE  
APPLICANT: WILKE, BETTINA  
APPLICANT: PREAC-MURSIG, VERA  
APPLICANT: MOTZ, MANFRED  
APPLICANT: SOUTSCHECK, ERWIN  
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS  
TITLE OF INVENTION: FROM BORRELLA BURGDOFFERI  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROOKS HAIDT HAEFNER & DELAHUNTY  
STREET: 99 PARK AVENUE  
CITY: NEW YORK  
STATE: NY  
COUNTRY: USA  
ZIP: 10016  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" FLOPPY DISC  
COMPUTER: AT&T - IBM COMPATIBLE  
OPERATING SYSTEM: MS-DOS Version 6.2  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/209,603E  
FILING DATE: 10-MAR-1994  
CLASSIFICATION: 436  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP90/02282  
FILING DATE: 21-DEC-1990  
APPLICATION NUMBER: US 07/862,535  
FILING DATE: 19-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: ROBINSON, WILLIAM R.  
REGISTRATION NUMBER: 27,224  
REFERENCE/DOCKET NUMBER: LGR-9217-A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 697-3355  
TELEFAX: (212) 557-5635  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 210  
TYPE: AMINO ACID

TOPOLOGY: LINEAR  
MOLECULE TYPE:  
DESCRIPTION: PROTEIN  
HYPOTHETICAL: N/A  
ANTI-SENSE: N/A  
FRAGMENT TYPE: N/A  
ORIGINAL SOURCE:  
ORGANISM: B. BURGDOFFERI  
IMMEDIATE SOURCE:  
LIBRARY: DSM 5662  
POSITION IN GENOME: N/A  
FEATURE:  
IDENTIFICATION METHOD: amino acid analysis  
PUBLICATION INFORMATION: N/A  
US-08-209-603E-15

Query Match 3.2%; Score 7; DB 4; Length 210;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 AKEALIK 45  
Db 155 AKEALIK 161

RESULT 5  
US-08-235-836C-30  
Sequence 30, Application US/08235836C  
Patent No. 6248562  
GENERAL INFORMATION:  
APPLICANT: Dunn, John J.  
APPLICANT: Luft, Benjamin J.  
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising  
TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor  
NUMBER OF SEQUENCES: 144  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Brookhaven National Laboratory  
STREET:  
CITY: Upton  
STATE: NY  
COUNTRY: USA  
ZIP: 11973  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/235,836C  
FILING DATE: 29-APR-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/148,191  
FILING DATE: 01-11-93  
ATTORNEY/AGENT INFORMATION:  
NAME: Bogosian, Margaret C.  
REGISTRATION NUMBER: 25,324  
REFERENCE/DOCKET NUMBER: ENL93-28A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 282-7338  
TELEFAX: (516) 282-3729  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 210 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-235-836C-30

Query Match 3.2%; Score 7; DB 4; Length 210;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 AKEALIK 45  
|||  
Db 155 AKEALIK 161

## RESULT 6

US-09-370-838-37  
; Sequence 37, Application US/09370838  
; Patent No. 6444425  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Mohamath, Roadoh  
; APPLICANT: Secrist, Heather  
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF  
; TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE  
; FILE REFERENCE: 210121.475C1  
; CURRENT FILING DATE: 1999-08-09  
; EARLIER FILING DATE: 1999-04-02  
; NUMBER OF SEQ ID NOS: 289  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 37  
; LENGTH: 238  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-370-838-37

Query Match 3.2%; Score 7; DB 4; Length 238;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 LNSGSVL 176  
|||  
Db 16 LNSGSVL 22

## RESULT 7

US-09-370-838-39  
; Sequence 39, Application US/09370838  
; Patent No. 6444425  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Mohamath, Roadoh  
; APPLICANT: Secrist, Heather  
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF  
; TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE  
; FILE REFERENCE: 210121.475C1  
; CURRENT FILING DATE: 1999-08-09  
; EARLIER FILING DATE: 1999-04-02  
; NUMBER OF SEQ ID NOS: 289  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 39  
; LENGTH: 243  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-370-838-39

Query Match 3.2%; Score 7; DB 4; Length 243;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 LNSGSVL 176  
|||  
Db 30 LNSGSVL 36

## RESULT 8

US-09-370-838-42

; Sequence 42, Application US/09370838  
; Patent No. 6444425  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Mohamath, Roadoh  
; APPLICANT: Secrist, Heather  
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF  
; TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE  
; FILE REFERENCE: 210121.475C1  
; CURRENT FILING DATE: 1999-08-09  
; EARLIER FILING DATE: 1999-04-02  
; NUMBER OF SEQ ID NOS: 289  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 42  
; LENGTH: 243  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-370-838-42

Query Match 3.2%; Score 7; DB 4; Length 243;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 LNSGSVL 176  
|||  
Db 21 LNSGSVL 27

## RESULT 9

US-09-370-838-43  
; Sequence 43, Application US/09370838  
; Patent No. 6444425  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Mohamath, Roadoh  
; APPLICANT: Secrist, Heather  
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF  
; TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE  
; FILE REFERENCE: 210121.475C1  
; CURRENT FILING DATE: 1999-08-09  
; EARLIER FILING DATE: 1999-04-02  
; NUMBER OF SEQ ID NOS: 289  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 43  
; LENGTH: 244  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-370-838-43

Query Match 3.2%; Score 7; DB 4; Length 244;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 LNSGSVL 176  
|||  
Db 22 LNSGSVL 28

## RESULT 10

US-09-370-838-46  
; Sequence 46, Application US/09370838  
; Patent No. 6444425  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Mohamath, Roadoh  
; APPLICANT: Secrist, Heather

```

; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.475C1
; CURRENT APPLICATION NUMBER: US/09/370,838
; EARLIER FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/285,323
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 46
; LENGTH: 244
; TYPE: PRN
; ORGANISM: Homo sapien
US-09-370-838-46

Query Match      3.2%; Score 7; DB 4; Length 244;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 LNSGSVL 176
DB 22 LNSGSVL 28

RESULT 11
US-09-370-838-40
; Sequence 40, Application US/09370838
; Patent No. 644425
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Roadoh
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.475C1
; CURRENT APPLICATION NUMBER: US/09/370,838
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/285,323
; EARLIER FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 40
; LENGTH: 245
; TYPE: PRN
; ORGANISM: Homo sapien
US-09-370-838-40

Query Match      3.2%; Score 7; DB 4; Length 245;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 LNSGSVL 176
DB 23 LNSGSVL 29

RESULT 12
US-08-235-836C-107
; Sequence 107, Application US/08235836C
; Patent No. 6248562
; GENERAL INFORMATION:
; APPLICANT: Dunn, John J.
; APPLICANT: Luft, Benjamin J.
; TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
; TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; STREET:
; ADDRESS: Brookhaven National Laboratory
; CITY: Upton
; STATE: NY
; COUNTRY: USA
```

```

; ZIP: 11973
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/235,836C
; FILING DATE: 29-Apr-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,191
; FILING DATE: 01-11-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Bogosian, Margaret C.
; REGISTRATION NUMBER: 25,324
; REFERENCE/DOCKET NUMBER: BNL93-28A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 282-3729
; TELEFAX: (516) 282-7338
; INFORMATION FOR SEQ ID NO: 107:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 466 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-235-836C-107

Query Match      3.2%; Score 7; DB 4; Length 466;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 AKEAILK 45
DB 411 AKEAILK 417

RESULT 13
US-08-235-836C-110
; Sequence 110, Application US/08235836C
; Patent No. 6248562
; GENERAL INFORMATION:
; APPLICANT: Dunn, John J.
; APPLICANT: Luft, Benjamin J.
; TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
; TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; STREET:
; ADDRESS: Brookhaven National Laboratory
; CITY: Upton
; STATE: NY
; COUNTRY: USA
; ZIP: 11973
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/235,836C
; FILING DATE: 29-Apr-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,191
; FILING DATE: 01-11-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Bogosian, Margaret C.
; REGISTRATION NUMBER: 25,324
; REFERENCE/DOCKET NUMBER: BNL93-28A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 282-3729
; TELEFAX: (516) 282-7338
```

INFORMATION FOR SEQ ID NO: 110:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 466 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-235-836C-110

Query Match 3.2%; Score 7; DB 4; Length 466;  
Best Local Similarity 100.0%; Pred. No. 41;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 AKEAILK 45  
Db 155 AKEAILK 161

## RESULT 14

US-08-235-836C-122  
Sequence 122, Application US/08235836C  
Patent No. 6248562  
GENERAL INFORMATION:  
APPLICANT: Dunn, John J.  
APPLICANT: Luft, Benjamin J.  
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising  
TITLE OF INVENTION: Borrelia Peptides and Uses Therefor  
NUMBER OF SEQUENCES: 144  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Brookhaven National Laboratory  
STREET:  
CITY: Upton  
STATE: NY  
COUNTRY: USA  
ZIP: 11973

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/235,836C  
FILING DATE: 29-APR-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/148,191  
FILING DATE: 01-11-93

ATTORNEY/AGENT INFORMATION:  
NAME: Bogosian, Margaret C.  
REGISTRATION NUMBER: 25,324  
REFERENCE/DOCKET NUMBER: BNT93-28A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 282-7338  
TELEFAX: (516) 282-3729  
INFORMATION FOR SEQ ID NO: 122:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 588 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-235-836C-122

Query Match 3.2%; Score 7; DB 4; Length 588;  
Best Local Similarity 100.0%; Pred. No. 50;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 AKEAILK 45  
Db 421 AKEAILK 427

RESULT 15  
US-09-562-737-95  
Sequence 95, Application US/09562737

Patent No. 6428967  
GENERAL INFORMATION:  
APPLICANT: Herz, Joachim  
APPLICANT: Gotthardt, Michael  
TITLE OF INVENTION: LDL Receptor Signaling Pathways  
FILE REFERENCE: UTSW0708  
CURRENT APPLICATION NUMBER: US/09/562,737  
CURRENT FILING DATE: 2000-05-01  
NUMBER OF SEQ ID NOS: 132  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 95  
LENGTH: 749  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-562-737-95

Query Match 3.2%; Score 7; DB 4; Length 749;  
Best Local Similarity 100.0%; Pred. No. 63;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 163 IIDQIVA 169  
Db 225 IIDQIVA 231

## RESULT 16

US-08-946-026-3  
Sequence 3, Application US/08946026  
Patent No. 6034218  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Twardzik, Daniel R.  
APPLICANT: Mitcham, Jennifer L.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
TITLE OF INVENTION: AND IMMUNODIAGNOSIS OF PROSTATE CANCER  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/946,026  
FILING DATE: 07-OCT-1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Makl, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.424C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 858 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-946-026-3

Query Match 3.2%; Score 7; DB 3; Length 858;  
Best Local Similarity 100.0%; Pred. No. 72;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 EGKAIID 165  
|||||  
Db 734 EGKAIID 740

RESULT 17  
US-09-206-942-41  
; Sequence 41, Application US/09206942  
; Patent No. 6432669  
; GENERAL INFORMATION:  
; APPLICANT: Loommore, Sheena M.  
; APPLICANT: Yang, Yan-Ping  
; APPLICANT: Klein, Michel H.  
; TITLE OF INVENTION: Protective Recombinant Haemophilus Influenzae High  
; FILE REFERENCE: 1038-861 MIS:jdb  
; CURRENT APPLICATION NUMBER: US/09/206,942  
; EARLIER FILING DATE: 1998-12-08  
; EARLIER APPLICATION NUMBER: 09/167,568  
; EARLIER FILING DATE: 1998-10-07  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 41  
; LENGTH: 1005  
; TYPE: PRT  
; ORGANISM: Haemophilus influenzae  
US-09-206-942-41

Query Match 3.2%; Score 7; DB 4; Length 1005;  
Best Local Similarity 100.0%; Pred. No. 83;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 VALNSGS 174  
|||||  
Db 231 VALNSGS 237

RESULT 18  
US-09-206-942-39  
; Sequence 39, Application US/09206942  
; Patent No. 6432669  
; GENERAL INFORMATION:  
; APPLICANT: Loommore, Sheena M.  
; APPLICANT: Yang, Yan-Ping  
; APPLICANT: Klein, Michel H.  
; TITLE OF INVENTION: Protective Recombinant Haemophilus Influenzae High  
; FILE REFERENCE: 1038-861 MIS:jdb  
; CURRENT APPLICATION NUMBER: US/09/206,942  
; EARLIER FILING DATE: 1998-12-08  
; EARLIER APPLICATION NUMBER: 09/167,568  
; EARLIER FILING DATE: 1998-10-07  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 39  
; LENGTH: 1011  
; TYPE: PRT  
; ORGANISM: Haemophilus influenzae  
US-09-206-942-39

Query Match 3.2%; Score 7; DB 4; Length 1011;  
Best Local Similarity 100.0%; Pred. No. 84;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 VALNSGS 174  
|||||  
Db 237 VALNSGS 243

RESULT 19  
US-09-162-564-6

; Sequence 6, Application US/09162564  
; Patent No. 6379664  
; GENERAL INFORMATION:  
; APPLICANT: Lou, Marjorie  
; APPLICANT: Raghavachari, Nalini  
; APPLICANT: Qiao, Fengyu  
; TITLE OF INVENTION: COMPOSITION AND METHOD FOR THE PREVENTION AND TREATMENT  
; FILE REFERENCE: UNTL-98-2-1  
; CURRENT APPLICATION NUMBER: US/09/162,564  
; CURRENT FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 107  
; TYPE: PRT  
; ORGANISM: Oryctolagus cuniculus  
US-09-162-564-6

Query Match 2.8%; Score 6; DB 4; Length 107;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 190 QEFVNS 195  
|||||  
Db 3 QEFVNS 8

RESULT 20  
US-08-207-481-31  
; Sequence 31, Application US/08207481  
; Patent No. 5820866  
; GENERAL INFORMATION:  
; APPLICANT: Kappler, John W.  
; APPLICANT: Mattrack, Philippa  
; TITLE OF INVENTION: PRODUCT AND PROCESS FOR T CELL  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: SHERIDAN ROSS & MCINTOSH  
; STREET: 1700 LINCOLN STREET, SUITE 3500  
; CITY: DENVER  
; STATE: COLORADO  
; COUNTRY: USA  
; ZIP: 80202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/207,481  
; FILING DATE: 04-MAR-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kovarik, Joseph E.  
; REGISTRATION NUMBER: 33,005  
; REFERENCE/DOCKET NUMBER: 2879-8  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 303/863-9700  
; TELEFAX: 303/863-0223  
; INFORMATION FOR SEQ ID NO: 31:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 121 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-207-481-31

Query Match 2.8%; Score 6; DB 2; Length 121;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 AAVILL 122  
| | | | |  
Db 11 AAVILL 16

RESULT 21  
PCT-US95-02689-33

; Sequence 33, Application PC/TUS9502689  
; GENERAL INFORMATION:  
; APPLICANT: National Jewish Center for Immunology and  
; APPLICANT: Respiratory Medicine  
; APPLICANT: Kapriel, John W.  
; APPLICANT: Marlick, Philippa  
; TITLE OF INVENTION: PRODUCT AND PROCESS FOR T CELL REGULATION  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SHERIDAN ROSS & MCINTOSH  
; STREET: 1700 LINCOLN STREET, SUITE 3500  
; CITY: DENVER  
; STATE: COLORADO  
; COUNTRY: USA  
; ZIP: 80202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/02689  
; FILING DATE: 03-MAR-1995  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kovarik, Joseph E.  
; REGISTRATION NUMBER: 33,005  
; REFERENCE/DOCKET NUMBER: 2879-8-PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 303/863-9700  
; TELEFAX: 303/863-0223  
; INFORMATION FOR SEQ ID NO: 33:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 121 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US95-02689-33

Query Match 2.8%; Score 6; DB 5; Length 121;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 AAVILL 122  
| | | | |  
Db 11 AAVILL 16

RESULT 22  
US-08-992-176-8

; Sequence 8, Application US/08992176  
; Patent No. 6125331  
; GENERAL INFORMATION:  
; APPLICANT: TOH, Hiroyuki  
; TITLE OF INVENTION: STRUCTURAL ALIGNMENT METHOD MAKING USE OF A DOUBLE  
; FILE REFERENCE: 9200-0001-2  
; CURRENT FILING DATE: 1997-12-17  
; EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 153  
; TYPE: PRT

; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: alpha-protein  
; OTHER INFORMATION: 1gdl)  
US-08-992-176-8

Query Match 2.8%; Score 6; DB 3; Length 153;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 KEALIK 45  
| | | | |  
Db 111 KEALIK 116

## RESULT 23

; US-09-040-213-2  
; Sequence 2, Application US/09040213  
; Patent No. 6287804  
; GENERAL INFORMATION:  
; APPLICANT: Black, Michael T.  
; TITLE OF INVENTION: nrdg  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dechert, Price & Rhoads  
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103-2793  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows 95  
; SOFTWARE: FastSeq for Windows Version 2.0b  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/040.213  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Falk, Stephen T  
; REGISTRATION NUMBER: 36,795  
; REFERENCE/DOCKET NUMBER: GM10136  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-994-2488  
; TELEFAX: 215-994-2222  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 178 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-09-040-213-2

Query Match 2.8%; Score 6; DB 4; Length 178;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 175 VLVDGL 180  
| | | | |  
Db 135 VLVDGL 140

RESULT 24  
US-09-026-958-2  
; Sequence 2, Application US/09026958  
; Patent No. 6150098  
; GENERAL INFORMATION:  
; APPLICANT: Zhang, Ke

APPLICANT: Pacific, Robert  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR IDENTIFYING  
TITLE OF INVENTION: NOVEL SECRETED MAMMALIAN POLYPEPTIDES  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Amgen Inc.  
STREET: One Amgen Center Drive  
CITY: Thousand Oaks  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 91320-1789  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/026,958  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Winter, Robert B.  
REFERENCE/DOCKET NUMBER: A-522  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 202 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-026-958-2

Query Match  
Best Local Similarity 2.8%; Score 6; DB 4; Length 202;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 PPOAAV 119  
Db 60 PPOAAV 65

RESULT 25  
US-09-057-860A-2  
Sequence 2, Application US/09057860A  
Patent No. 6277820  
GENERAL INFORMATION:  
APPLICANT: Arnon Rosenthal  
APPLICANT: Mary Hynes  
TITLE OF INVENTION: Method Of Dopaminergic And Serotonergic  
TITLE OF INVENTION: Neuron Formation From Neuroprogenitor Cells  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Winpatin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/057,860A  
FILING DATE: 09-Apr-1998  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Svoboda, Craig G.  
REGISTRATION NUMBER: 39,044  
REFERENCE/DOCKET NUMBER: P1364  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-1489

TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 202 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
US-09-057-860A-2

Query Match  
Best Local Similarity 2.8%; Score 6; DB 4; Length 202;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 PPOAAV 119  
Db 60 PPOAAV 65

RESULT 26  
US-08-177-109A-60  
Sequence 60, Application US/08177109A  
Patent No. 5869615  
GENERAL INFORMATION:  
APPLICANT: Dennis E. Hourcade and Teresa J. Oglesby  
TITLE OF INVENTION: MODIFIED COMPLEMENT PROTEASES  
NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Patrea L. Pabst  
STREET: 2801 One Atlantic Center  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: USA  
ZIP: 30309-3450  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/177,109A  
FILING DATE: 03-JAN-1994  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: WU 107  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (404) 873-8794  
TELEFAX: (404) 873-8795  
INFORMATION FOR SEQ ID NO: 60:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 207 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
US-08-177-109A-60

Query Match  
Best Local Similarity 2.8%; Score 6; DB 2; Length 207;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 ILITDG 125  
Db 125 ILITDG 130

RESULT 27  
US-08-687-706-60  
Sequence 60, Application US/08687706  
Patent No. 5928892  
GENERAL INFORMATION:  
APPLICANT: Dennis E. Hourcade and Teresa J. Oglesby



TITLE OF INVENTION: MODIFIED COMPLEMENT PROTEASES  
NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Patrea L. Pabst  
STREET: 2800 One Atlantic Center  
STREET: 1201 West Peachtree Street  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: USA  
ZIP: 30309-3450  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/687,706  
FILING DATE: 26-JUL-1996  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/177,109  
FILING DATE: 03-JAN-1994  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: MU 107 DIV  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (404) 873-8794  
TELEFAX: (404) 873-8795  
INFORMATION FOR SEQ ID NO: 60:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 207 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
US-08-687-706-60

Query Match 2.8%; Score 6; DB 2; Length 207;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 120 ILITDG 125  
Db 125 ILITDG 130

RESULT 28  
US-08-031-295-2  
Sequence 2, Application US/08031295  
Patent No. 5510103  
GENERAL INFORMATION:  
APPLICANT: LIVEY, Ian  
APPLICANT: DORNER, Friedrich  
TITLE OF INVENTION: METHOD AND COMPOSITION FOR THE  
PREVENTION OF LYME DISEASE  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington, D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/031,295  
FILING DATE: 19930312  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/903,580  
FILING DATE: 25-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/824,161  
FILING DATE: 22-JAN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/727,245  
FILING DATE: 11-JUL-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/142 IMMU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 672-5300  
TELEFAX: (202) 672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 212 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-031-295-2

Query Match 2.8%; Score 6; DB 1; Length 212;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 40 KEALIK 45  
Db 158 KEALIK 163

RESULT 29  
US-07-903-580-2  
Sequence 2, Application US/07903580  
Patent No. 6221363  
GENERAL INFORMATION:  
APPLICANT: LIVEY, Ian  
APPLICANT: DORNER, Friedrich  
TITLE OF INVENTION: METHOD AND COMPOSITION FOR THE  
PREVENTION OF LYME DISEASE  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/903,580  
FILING DATE: 19920625  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/824,161  
FILING DATE: 22-JAN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/727,245  
FILING DATE: 11-JUL-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/131 IMMU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-9300  
TELEFAX: (703) 683-4109

```

;
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 212 amino acids
;   TYPE: AMINO ACID
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-903-580-2

Query Match          2.8%; Score 6; DB 4; Length 212;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 KEALIK 45
Db 158 KEALIK 163

RESULT 30
US-09-134-001C-2945
; Sequence 2945, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2945
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-2945

Query Match          2.8%; Score 6; DB 4; Length 226;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 PVEEVK 138
Db 62 PVEEVK 67

RESULT 31
US-08-512-955-4
; Sequence 4, Application US/08512955
; Patent No. 5976536
; GENERAL INFORMATION:
; APPLICANT: Stephens, David S.
; APPLICANT: Kahler, Charlene M.
; TITLE OF INVENTION: Neisseria Mutants, Lipooligosaccharides
; TITLE OF INVENTION: and Immunogenic Compositions
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee and Winner, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/512,955
```

```

;
; FILING DATE: 09-AUG-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Feberer, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 12-95
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 242 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-512-955-4

Query Match          2.8%; Score 6; DB 2; Length 242;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 VALNSG 173
Db 177 VALNSG 182

RESULT 32
US-09-627-376-14
; Sequence 14, Application US/09627376
; Patent No. 6342385
; GENERAL INFORMATION:
; APPLICANT: Qi, Fengxia
; TITLE OF INVENTION: MUTACIN I BIOSYNTHESIS GENES AND PROTEINS
; FILE REFERENCE: UAB-17402/22
; CURRENT APPLICATION NUMBER: US/09/627,376
; CURRENT FILING DATE: 2001-05-30
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Streptococcus mutans
; US-09-627-376-14

Query Match          2.8%; Score 6; DB 4; Length 246;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 AVILLT 123
Db 132 AVILLT 137

RESULT 33
US-09-452-239-20
; Sequence 20, Application US/09452239
; Patent No. 6465229
; GENERAL INFORMATION:
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Fader, Gary M.
; APPLICANT: Caroon, Rebecca E.
; TITLE OF INVENTION: Plant Caffeoyl-CoA O-Methyltransferase
; FILE REFERENCE: BR1284 US NA
; CURRENT APPLICATION NUMBER: US/09/452,239
; CURRENT FILING DATE: 1999-12-01
; EARLIER APPLICATION NUMBER: 60/110,594
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 20
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Glycine max
```

US-09-452-239-20

Query Match 2.8%; Score 6; DB 4; Length 246;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 98 PVGDGI 103  
|||||  
DB 234 PVGDGI 239

RESULT 34  
US-09-452-239-22  
; Sequence 22, Application US/09452239  
; Patent No. 6465229  
; GENERAL INFORMATION:  
; APPLICANT: Rafalski, Antoni J.  
; APPLICANT: Fader, Gary M.  
; APPLICANT: Cahoon, Rebecca E.  
; TITLE OF INVENTION: Plant Caffeoyl-CoA O-Methyltransferase  
; FILE REFERENCE: BB1284 US NA  
; CURRENT APPLICATION NUMBER: US/09/452,239  
; CURRENT FILING DATE: 1999-12-01  
; EARLIER APPLICATION NUMBER: 60/110,594  
; EARLIER FILING DATE: 1998-December-02  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 22  
; LENGTH: 246  
; TYPE: PRT  
; ORGANISM: Glycine max  
US-09-452-239-22

Query Match 2.8%; Score 6; DB 4; Length 246;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 98 PVGDGI 103  
|||||  
DB 234 PVGDGI 239

RESULT 35  
US-09-452-239-18  
; Sequence 18, Application US/09452239  
; Patent No. 6465229  
; GENERAL INFORMATION:  
; APPLICANT: Rafalski, Antoni J.  
; APPLICANT: Fader, Gary M.  
; APPLICANT: Cahoon, Rebecca E.  
; TITLE OF INVENTION: Plant Caffeoyl-CoA O-Methyltransferase  
; FILE REFERENCE: BB1284 US NA  
; CURRENT APPLICATION NUMBER: US/09/452,239  
; CURRENT FILING DATE: 1999-12-01  
; EARLIER APPLICATION NUMBER: 60/110,594  
; EARLIER FILING DATE: 1998-December-02  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 18  
; LENGTH: 247  
; TYPE: PRT  
; ORGANISM: Glycine max  
US-09-452-239-18

Query Match 2.8%; Score 6; DB 4; Length 247;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 98 PVGDGI 103  
|||||  
DB 235 PVGDGI 240

RESULT 36  
US-09-452-239-50  
; Sequence 50, Application US/09452239  
; Patent No. 6465229  
; GENERAL INFORMATION:  
; APPLICANT: Rafalski, Antoni J.  
; APPLICANT: Fader, Gary M.  
; APPLICANT: Cahoon, Rebecca E.  
; TITLE OF INVENTION: Plant Caffeoyl-CoA O-Methyltransferase  
; FILE REFERENCE: BB1284 US NA  
; CURRENT APPLICATION NUMBER: US/09/452,239  
; CURRENT FILING DATE: 1999-12-01  
; EARLIER APPLICATION NUMBER: 60/110,594  
; EARLIER FILING DATE: 1998-December-02  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 50  
; LENGTH: 247  
; TYPE: PRT  
; ORGANISM: Populus tremuloides  
US-09-452-239-50

Query Match 2.8%; Score 6; DB 4; Length 247;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 98 PVGDGI 103  
|||||  
DB 235 PVGDGI 240

RESULT 37  
US-09-189-0608-12  
; Sequence 12, Application US/091890608  
; Patent No. 6270968  
; GENERAL INFORMATION:  
; APPLICANT: Dalboge, Henrik  
; APPLICANT: Sanddal, Thomas  
; APPLICANT: Kauppinen, Markus  
; APPLICANT: Borge, Diderichsen  
; TITLE OF INVENTION: Method of Providing No. 6270968el DNA Sequences  
; FILE REFERENCE: 4772.204-US  
; CURRENT APPLICATION NUMBER: US/09/189,0608  
; CURRENT FILING DATE: 1998-11-10  
; PRIOR APPLICATION NUMBER: PCT/DK97/00216  
; PRIOR FILING DATE: 1997-05-12  
; NUMBER OF SEQ ID NOS: 74  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12  
; LENGTH: 248  
; TYPE: PRT  
; ORGANISM: Hybrid  
US-09-189-0608-12

Query Match 2.8%; Score 6; DB 4; Length 248;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 201 QILVTE 206  
|||||  
DB 206 QILVTE 211

RESULT 38  
US-09-452-239-24  
; Sequence 24, Application US/09452239  
; Patent No. 6465229  
; GENERAL INFORMATION:  
; APPLICANT: Rafalski, Antoni J.  
; APPLICANT: Fader, Gary M.  
; APPLICANT: Cahoon, Rebecca E.  
; TITLE OF INVENTION: Plant Caffeoyl-CoA O-Methyltransferase  
; FILE REFERENCE: BB1284 US NA

; CURRENT APPLICATION NUMBER: US/09/452,239  
 ; CURRENT FILING DATE: 1999-12-01  
 ; EARLIER APPLICATION NUMBER: 60/110,594  
 ; EARLIER FILING DATE: 1998-December-02  
 ; NUMBER OF SEQ ID NOS: 50  
 ; SOFTWARE: Microsoft Office 97  
 ; SEQ ID NO 24  
 ; LENGTH: 248  
 ; TYPE: PRT  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; NAME/KEY: UNSURE  
 ; LOCATION: (7)  
 ; US-09-452-239-24

Query Match 2.8%; Score 6; DB 4; Length 248;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 PVGDGI 103  
 |||||  
 Db 236 PVGDGI 241

RESULT 39  
 ; US-09-452-239-26  
 ; Sequence 26, Application US/09452239  
 ; Patent No. 6465229  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rafaleki, Antoni J.  
 ; APPLICANT: Fader, Gary M.  
 ; APPLICANT: Cahoon, Rebecca E.  
 ; TITLE OF INVENTION: Plant Caffeoyl-CoA O-Methyltransferase  
 ; FILE REFERENCE: BB1284 US NA  
 ; CURRENT APPLICATION NUMBER: US/09/452,239  
 ; CURRENT FILING DATE: 1999-12-01  
 ; EARLIER APPLICATION NUMBER: 60/110,594  
 ; EARLIER FILING DATE: 1998-December-02  
 ; NUMBER OF SEQ ID NOS: 50  
 ; SOFTWARE: Microsoft Office 97  
 ; SEQ ID NO 26  
 ; LENGTH: 248  
 ; TYPE: PRT  
 ; ORGANISM: Glycine max  
 ; US-09-452-239-26

Query Match 2.8%; Score 6; DB 4; Length 248;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 PVGDGI 103  
 |||||  
 Db 236 PVGDGI 241

RESULT 40  
 ; US-08-207-481-20  
 ; Sequence 20, Application US/08207481  
 ; Patent No. 5820866  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kappeler, John W.  
 ; APPLICANT: Marrack, Philippa  
 ; TITLE OF INVENTION: PRODUCT AND PROCESS FOR T CELL  
 ; TITLE OF INVENTION: REGULATION  
 ; NUMBER OF SEQUENCES: 45  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SHERIDAN ROSS & MCINTOSH  
 ; STREET: 1700 LINCOLN STREET, SUITE 3500  
 ; CITY: DENVER  
 ; STATE: COLORADO  
 ; COUNTRY: USA  
 ; ZIP: 80202  
 ; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/207,481  
 ; FILING DATE: 04-MAR-1994  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Kovarik, Joseph E.  
 ; REGISTRATION NUMBER: 33,005  
 ; REFERENCE/DOCKET NUMBER: 2879-8  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 303/863-9700  
 ; TELEFAX: 303/863-0223  
 ; INFORMATION FOR SEQ ID NO: 20:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 254 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-207-481-20

Query Match 2.8%; Score 6; DB 2; Length 254;  
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 AAVTLL 122  
 |||||  
 Db 11 AAVTLL 16

Search completed: April 9, 2003, 14:46:24  
 Job time : 43 secs



GenCore version 5.1.4 p5\_4578  
Copyright (c) 1993 - 2003 Comugen Ltd.

## OM protein - protein search, using sw model

Run on: April 9, 2003, 14:45:29 ; Search time 36 Seconds  
(without alignments)  
370.213 Million cell updates/sec

Title: US-10-034-500-2

Perfect score: 218

Sequence: 1 AEVTASCTKRVESYNYLVLDY.....QEOILVTEVVVLGVNFAP 218

Scoring table: OLIGO

Searched: 248812 seqs, 61136040 residues

Word size : 0

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

## Database :

Published Applications AA:\*  
1: /cgn2\_6/prodata/2/pubppaa/US08\_NEW\_PUB pep:\*  
2: /cgn2\_6/prodata/2/pubppaa/PCT\_NEW\_PUB pep:\*  
3: /cgn2\_6/prodata/2/pubppaa/US06\_NEW\_PUB pep:\*  
4: /cgn2\_6/prodata/2/pubppaa/US06\_PUBCOMB pep:\*  
5: /cgn2\_6/prodata/2/pubppaa/US07\_NEW\_PUB pep:\*  
6: /cgn2\_6/prodata/2/pubppaa/US07\_PUBCOMB pep:\*  
7: /cgn2\_6/prodata/2/pubppaa/PCTUS\_PUBCOMB pep:\*  
8: /cgn2\_6/prodata/2/pubppaa/US08\_PUBCOMB pep:\*  
9: /cgn2\_6/prodata/2/pubppaa/US09\_NEW\_PUB pep:\*  
10: /cgn2\_6/prodata/2/pubppaa/US09\_PUBCOMB pep:\*  
11: /cgn2\_6/prodata/2/pubppaa/US10\_NEW\_PUB pep:\*  
12: /cgn2\_6/prodata/2/pubppaa/US10\_PUBCOMB pep:\*  
13: /cgn2\_6/prodata/2/pubppaa/US60\_NEW\_PUB pep:\*  
14: /cgn2\_6/prodata/2/pubppaa/US60\_PUBCOMB pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	3.2	151	9 US-10-101-464A-551	Sequence 551, App
2	7	3.2	238	9 US-09-854-133-37	Sequence 37, App
3	7	3.2	238	10 US-09-738-973-37	Sequence 37, App
4	7	3.2	243	9 US-09-854-133-39	Sequence 39, App
5	7	3.2	243	9 US-09-854-133-42	Sequence 42, App
6	7	3.2	243	10 US-09-738-973-39	Sequence 39, App
7	7	3.2	243	10 US-09-738-973-42	Sequence 42, App
8	7	3.2	244	9 US-09-854-133-43	Sequence 43, App
9	7	3.2	244	9 US-09-854-133-46	Sequence 46, App
10	7	3.2	244	10 US-09-738-973-43	Sequence 43, App
11	7	3.2	244	10 US-09-738-973-46	Sequence 46, App
12	7	3.2	245	9 US-09-854-133-40	Sequence 40, App
13	7	3.2	245	10 US-09-738-973-40	Sequence 40, App
14	7	3.2	311	10 US-09-815-242-11873	Sequence 11873, A
15	6	2.8	14	10 US-09-773-882-4	Sequence 4, App
16	6	2.8	15	10 US-09-976-674-61	Sequence 61, App
17	6	2.8	28	10 US-09-864-761-41312	Sequence 41312, A
18	6	2.8	34	10 US-09-864-761-36602	Sequence 36602, A
19	6	2.8	43	10 US-09-864-761-33862	Sequence 33862, A

20	6	2.8	58	9 US-09-981-676-242	Sequence 242, App
21	6	2.8	58	9 US-09-148-545-242	Sequence 242, App
22	6	2.8	59	9 US-09-981-876-186	Sequence 186, App
23	6	2.8	59	9 US-09-148-545-186	Sequence 186, App
24	6	2.8	62	10 US-09-864-761-45365	Sequence 45365, A
25	6	2.8	70	10 US-09-864-761-46027	Sequence 46027, A
26	6	2.8	87	9 US-09-738-626-674	Sequence 674, App
27	6	2.8	88	9 US-09-984-271-178	Sequence 178, App
28	6	2.8	88	9 US-09-984-271-179	Sequence 179, App
29	6	2.8	93	10 US-09-864-761-37228	Sequence 37228, A
30	6	2.8	100	9 US-10-012-896-933	Sequence 933, App
31	6	2.8	100	9 US-09-895-793-933	Sequence 933, App
32	6	2.8	100	9 US-09-895-814-933	Sequence 933, App
33	6	2.8	100	10 US-09-759-143-933	Sequence 933, App
34	6	2.8	100	10 US-09-780-669-933	Sequence 933, App
35	6	2.8	100	10 US-09-822-827-933	Sequence 933, App
36	6	2.8	109	9 US-09-786-592-1175	Sequence 1175, App
37	6	2.8	122	9 US-09-957-708-10	Sequence 30, App
38	6	2.8	132	9 US-09-738-626-5132	Sequence 5132, App
39	6	2.8	139	10 US-09-764-853-560	Sequence 560, App
40	6	2.8	150	10 US-09-815-242-13435	Sequence 13435, A
41	6	2.8	154	9 US-10-260-877-4	Sequence 4, App
42	6	2.8	177	9 US-09-965-528-17	Sequence 17, App
43	6	2.8	178	10 US-09-884-519-2	Sequence 2, App
44	6	2.8	178	10 US-09-884-441-489	Sequence 489, App
45	6	2.8	181	9 US-09-984-271-114	Sequence 114, App
46	6	2.8	187	10 US-09-242-980-3	Sequence 3, App
47	6	2.8	188	9 US-09-986-480-258	Sequence 258, App
48	6	2.8	193	9 US-10-172-088-11	Sequence 11, App
49	6	2.8	197	9 US-10-043-487-331	Sequence 31, App
50	6	2.8	208	9 US-09-895-913A-360	Sequence 360, App
51	6	2.8	209	9 US-10-012-896-897	Sequence 897, App
52	6	2.8	209	9 US-09-895-793-897	Sequence 897, App
53	6	2.8	209	9 US-09-895-814-897	Sequence 897, App
54	6	2.8	209	10 US-09-759-143-897	Sequence 897, App
55	6	2.8	209	10 US-09-780-669-897	Sequence 897, App
56	6	2.8	209	10 US-09-822-827-897	Sequence 897, App
57	6	2.8	213	10 US-09-844-468-4	Sequence 4, App
58	6	2.8	220	10 US-09-864-761-42927	Sequence 42927, A
59	6	2.8	222	9 US-09-895-913A-278	Sequence 278, App
60	6	2.8	223	9 US-10-002-050-4	Sequence 4, App
61	6	2.8	223	9 US-10-002-304-4	Sequence 4, App
62	6	2.8	223	12 US-10-003-152-4	Sequence 4, App
63	6	2.8	224	10 US-09-242-980-13	Sequence 13, App
64	6	2.8	233	9 US-09-738-626-5648	Sequence 5648, App
65	6	2.8	233	10 US-09-884-441-488	Sequence 488, App
66	6	2.8	240	12 US-10-001-843-197	Sequence 197, App
67	6	2.8	246	10 US-09-452-239-22	Sequence 22, App
68	6	2.8	246	10 US-09-452-239-22	Sequence 22, App
69	6	2.8	246	12 US-10-047-676A-14	Sequence 14, App
70	6	2.8	247	10 US-09-452-239-18	Sequence 18, App
71	6	2.8	247	10 US-09-452-239-50	Sequence 50, App
72	6	2.8	248	10 US-09-452-239-24	Sequence 24, App
73	6	2.8	248	10 US-09-452-239-26	Sequence 26, App
74	6	2.8	258	9 US-10-023-437-29	Sequence 29, App
75	6	2.8	259	10 US-09-452-239-42	Sequence 42, App
76	6	2.8	260	10 US-09-452-239-12	Sequence 12, App
77	6	2.8	262	10 US-09-860-738A-2	Sequence 2, App
78	6	2.8	263	10 US-09-738-626-3560	Sequence 3560, App
79	6	2.8	264	10 US-09-990-017-2	Sequence 2, App
80	6	2.8	271	9 US-09-308-007-13	Sequence 33, App
81	6	2.8	276	9 US-09-738-626-5599	Sequence 5599, App
82	6	2.8	277	9 US-09-766-753-154	Sequence 154, App
83	6	2.8	283	10 US-09-828-313-30	Sequence 30, App
84	6	2.8	286	9 US-09-738-626-8848	Sequence 8848, App
85	6	2.8	294	10 US-09-927-738-9	Sequence 9, App
86	6	2.8	296	9 US-09-796-753-152	Sequence 152, App
87	6	2.8	312	10 US-09-815-242-11339	Sequence 11339, App
88	6	2.8	312	10 US-09-815-242-11511	Sequence 11511, App
89	6	2.8	312	10 US-09-935-428A-4	Sequence 4, App
90	6	2.8	315	9 US-10-104-966-5	Sequence 5, App
91	6	2.8	315	10 US-09-929-955-5	Sequence 5, App
92	6	2.8	330	9 US-09-984-245-203	Sequence 203, App

93	6	2.8	330	9	US-09-966-262-203	Sequence 203, App	166	6	2.8	638	9	US-09-811-469-9	Sequence 9, App1
94	6	2.8	330	9	US-09-983-966-203	Sequence 203, App	167	6	2.8	659	9	US-10-227-490-11	Sequence 11, App1
95	6	2.8	345	10	US-09-815-242-5354	Sequence 5354, App	168	6	2.8	662	10	US-09-815-242-11390	Sequence 11390, A
96	6	2.8	347	9	US-09-738-626-6559	Sequence 6559, App	169	6	2.8	663	9	US-10-108-605-245	Sequence 245, App
97	6	2.8	356	9	US-10-084-205-62	Sequence 62, App1	170	6	2.8	677	10	US-09-815-242-12100	Sequence 12100, A
98	6	2.8	356	10	US-09-925-637-62	Sequence 62, App1	171	6	2.8	684	9	US-10-097-534-11	Sequence 11, App1
99	6	2.8	360	9	US-09-372-348-9	Sequence 9, App1	172	6	2.8	691	10	US-09-976-674-43	Sequence 43, App1
100	6	2.8	370	10	US-09-864-761-35913	Sequence 35913, A	173	6	2.8	706	10	US-10-097-534-10	Sequence 10, App1
101	6	2.8	374	10	US-09-815-242-13473	Sequence 13473, A	174	6	2.8	739	9	US-10-097-534-10	Sequence 10, App1
102	6	2.8	374	10	US-09-815-242-13650	Sequence 13650, A	175	6	2.8	744	9	US-10-096-961-4	Sequence 4, App1
103	6	2.8	387	9	US-09-712-363-256	Sequence 256, App	176	6	2.8	744	9	US-10-096-961-5	Sequence 5, App1
104	6	2.8	393	9	US-09-813-398-34	Sequence 34, App1	177	6	2.8	750	9	US-10-227-490-13	Sequence 13, App1
105	6	2.8	393	10	US-09-815-242-12395	Sequence 12395, A	178	6	2.8	750	9	US-10-227-490-14	Sequence 14, App1
106	6	2.8	414	10	US-09-884-441-483	Sequence 706, App	179	6	2.8	795	9	US-10-055-4128-11	Sequence 11, App1
107	6	2.8	414	10	US-09-925-302-706	Sequence 8, App1	180	6	2.8	796	9	US-10-160-501-5	Sequence 5, App1
108	6	2.8	417	9	US-09-935-642-7	Sequence 7, App1	181	6	2.8	796	10	US-09-976-674-5	Sequence 5, App1
109	6	2.8	432	9	US-09-738-626-3771	Sequence 3771, App	182	6	2.8	821	9	US-10-055-4128-12	Sequence 12, App1
110	6	2.8	436	10	US-09-925-301-928	Sequence 928, App	183	6	2.8	833	10	US-09-884-441-389	Sequence 389, App
111	6	2.8	438	10	US-09-884-441-483	Sequence 390, App	184	6	2.8	839	9	US-10-227-490-9	Sequence 9, App1
112	6	2.8	438	10	US-09-884-441-483	Sequence 483, App	185	6	2.8	839	9	US-10-227-490-10	Sequence 10, App1
113	6	2.8	446	9	US-09-975-719-267	Sequence 267, App	186	6	2.8	841	9	US-09-533-029-102	Sequence 102, App
114	6	2.8	447	9	US-09-738-626-6982	Sequence 6982, App	187	6	2.8	903	9	US-10-055-4128-46	Sequence 46, App1
115	6	2.8	464	12	US-10-095-492-14	Sequence 14, App1	188	6	2.8	905	9	US-10-055-4128-2	Sequence 2, App1
116	6	2.8	465	10	US-09-923-922-6	Sequence 6, App1	189	6	2.8	911	10	US-09-745-163-140	Sequence 140, App
117	6	2.8	465	10	US-09-923-922-10	Sequence 10, App1	190	6	2.8	914	10	US-09-778-320-206	Sequence 206, App
118	6	2.8	474	9	US-09-774-639-371	Sequence 371, App	191	6	2.8	914	10	US-09-910-689-206	Sequence 206, App
119	6	2.8	474	9	US-09-969-730-249	Sequence 249, App	192	6	2.8	914	10	US-09-884-441-312	Sequence 312, App
120	6	2.8	475	10	US-09-741-659-458	Sequence 458, App	193	6	2.8	914	10	US-09-884-441-478	Sequence 478, App
121	6	2.8	484	9	US-09-738-626-6430	Sequence 6430, App	194	6	2.8	914	12	US-10-010-742-206	Sequence 206, App
122	6	2.8	484	9	US-10-012-896-895	Sequence 895, App	195	6	2.8	931	10	US-09-906-453-12	Sequence 12, App1
123	6	2.8	492	9	US-10-012-896-932	Sequence 932, App	196	6	2.8	969	9	US-09-961-353-122	Sequence 122, App
124	6	2.8	492	9	US-09-895-793-895	Sequence 895, App	197	6	2.8	969	9	US-10-023-437-33	Sequence 33, App1
125	6	2.8	492	9	US-09-895-793-932	Sequence 932, App	198	6	2.8	977	10	US-09-925-297-797	Sequence 797, App
126	6	2.8	492	9	US-09-895-814-895	Sequence 895, App	199	6	2.8	997	10	US-09-747-771-3	Sequence 3, App1
127	6	2.8	492	9	US-09-895-814-932	Sequence 932, App	200	6	2.8	999	10	US-09-747-771-2	Sequence 2, App1
128	6	2.8	492	10	US-09-759-143-895	Sequence 895, App	201	6	2.8	1000	9	US-10-055-4128-30	Sequence 30, App1
129	6	2.8	492	10	US-09-759-143-932	Sequence 932, App	202	6	2.8	1028	9	US-09-947-063-13	Sequence 13, App1
130	6	2.8	492	10	US-09-780-669-895	Sequence 895, App	203	6	2.8	1094	9	US-10-043-487-300	Sequence 300, App
131	6	2.8	492	10	US-09-780-669-932	Sequence 932, App	204	6	2.8	1119	10	US-09-782-714-2	Sequence 2, App1
132	6	2.8	492	10	US-09-879-792-14	Sequence 14, App1	205	6	2.8	1148	10	US-09-884-441-458	Sequence 458, App
133	6	2.8	492	10	US-09-822-827-895	Sequence 895, App	206	6	2.8	1148	10	US-09-884-441-479	Sequence 479, App
134	6	2.8	492	10	US-09-822-827-932	Sequence 932, App	207	6	2.8	1156	10	US-09-884-441-453	Sequence 459, App
135	6	2.8	497	10	US-09-892-325-4	Sequence 4, App1	208	6	2.8	1170	12	US-10-135-687-2	Sequence 2, App1
136	6	2.8	500	10	US-09-883-797-12	Sequence 12, App1	209	6	2.8	1185	9	US-09-895-913A-246	Sequence 246, App
137	6	2.8	506	10	US-09-877-476-2	Sequence 2, App1	210	6	2.8	1185	10	US-09-815-242-11466	Sequence 11466, A
138	6	2.8	506	10	US-09-877-476-8	Sequence 8, App1	211	6	2.8	1192	10	US-09-789-386-2	Sequence 2, App1
139	6	2.8	506	10	US-09-877-476-12	Sequence 12, App1	212	6	2.8	1192	10	US-09-758-140-6	Sequence 6, App1
140	6	2.8	506	10	US-09-877-476-14	Sequence 14, App1	213	6	2.8	1192	10	US-09-893-348-23	Sequence 23, App1
141	6	2.8	506	10	US-09-877-476-16	Sequence 16, App1	214	6	2.8	1192	10	US-09-972-599A-6	Sequence 6, App1
142	6	2.8	506	10	US-09-877-476-20	Sequence 20, App1	215	6	2.8	1210	10	US-09-860-352A-2	Sequence 2, App1
143	6	2.8	506	10	US-09-877-476-32	Sequence 22, App1	216	6	2.8	1319	9	US-10-041-856-8	Sequence 8, App1
144	6	2.8	506	10	US-09-877-476-34	Sequence 24, App1	217	6	2.8	1488	9	US-10-043-487-385	Sequence 285, App
145	6	2.8	506	10	US-09-877-476-26	Sequence 26, App1	218	6	2.8	1945	9	US-09-927-587-2	Sequence 2, App1
146	6	2.8	506	10	US-09-877-476-30	Sequence 30, App1	219	6	2.8	1979	9	US-09-927-587-4	Sequence 4, App1
147	6	2.8	506	10	US-09-877-476-36	Sequence 36, App1	220	6	2.8	2201	12	US-10-029-907-3	Sequence 3, App1
148	6	2.8	506	10	US-09-877-476-38	Sequence 38, App1	221	6	2.8	2383	10	US-09-912-020-302	Sequence 302, App
149	6	2.8	506	10	US-09-877-476-40	Sequence 40, App1	222	6	2.8	3011	9	US-10-104-966-1	Sequence 1, App1
150	6	2.8	507	9	US-09-962-678-4	Sequence 4, App1	223	6	2.8	3011	10	US-09-899-855-1	Sequence 1, App1
151	6	2.8	507	10	US-09-895-728-4	Sequence 4, App1	224	6	2.8	3298	9	US-10-149-819-21	Sequence 21, App1
152	6	2.8	510	9	US-09-738-626-6798	Sequence 6798, App	225	6	2.8	3916	10	US-09-737-149-29	Sequence 29, App1
153	6	2.8	511	10	US-09-773-882-2	Sequence 2, App1	226	6	2.8	29326	9	US-09-759-508B-2	Sequence 2, App1
154	6	2.8	511	10	US-09-773-882-11	Sequence 11, App1	227	5	2.3	8	10	US-09-924-824-72	Sequence 72, App
155	6	2.8	512	10	US-09-815-242-11232	Sequence 11232, A	228	5	2.3	10	9	US-09-826-280-202	Sequence 202, App
156	6	2.8	517	10	US-09-801-368-100	Sequence 100, App	229	5	2.3	13	9	US-10-099-135-22	Sequence 22, App1
157	6	2.8	518	10	US-09-815-242-12473	Sequence 12473, A	230	5	2.3	15	10	US-09-791-178-465	Sequence 465, App
158	6	2.8	530	9	US-10-107-695B-2	Sequence 2, App1	231	5	2.3	16	10	US-09-826-1752-18	Sequence 18, App1
159	6	2.8	530	9	US-09-372-348-11	Sequence 11, App1	232	5	2.3	18	10	US-09-839-497A-3	Sequence 3, App1
160	6	2.8	537	10	US-09-883-797-10	Sequence 10, App1	233	5	2.3	232	10	US-10-036-492-7	Sequence 7, App1
161	6	2.8	559	10	US-09-815-242-12086	Sequence 12086, A	234	5	2.3	18	10	US-09-864-761-40757	Sequence 40757, A
162	6	2.8	570	9	US-10-227-490-15	Sequence 15, App1	235	5	2.3	20	9	US-10-079-478-14	Sequence 14, App1
163	6	2.8	590	9	US-09-738-626-3983	Sequence 3983, App	236	5	2.3	20	9	US-10-079-478-15	Sequence 15, App1
164	6	2.8	600	10	US-09-801-368-158	Sequence 158, App	237	5	2.3	20	10	US-09-864-761-47919	Sequence 47919, A
165	6	2.8	612	10	US-09-758-269-12	Sequence 12, App1	238	5	2.3	21	9	US-09-962-445-6	Sequence 6, App1

239	5	2.3	21	9	US-09-974-879-410	Sequence 410, App	312	5	2.3	39	10	US-09-864-761-37706	Sequence 37706, A
240	5	2.3	21	10	US-09-853-830-2	Sequence 2, App1	313	5	2.3	39	10	US-09-764-847-937	Sequence 937, App
241	5	2.3	21	10	US-09-853-830-3	Sequence 3, App1	314	5	2.3	40	9	US-10-042-141-94	Sequence 94, App1
242	5	2.3	21	10	US-09-853-830-4	Sequence 4, App1	315	5	2.3	40	10	US-09-726-643-94	Sequence 94, App1
243	5	2.3	21	10	US-09-853-830-5	Sequence 5, App1	316	5	2.3	42	9	US-09-776-724-169	Sequence 169, App
244	5	2.3	21	10	US-09-853-830-6	Sequence 6, App1	317	5	2.3	42	10	US-09-864-761-34791	Sequence 34791, A
245	5	2.3	21	10	US-09-853-830-7	Sequence 7, App1	318	5	2.3	43	9	US-09-861-876-205	Sequence 205, App
246	5	2.3	21	10	US-09-853-830-8	Sequence 8, App1	319	5	2.3	43	9	US-09-939-780-3	Sequence 3, App1
247	5	2.3	21	10	US-09-853-830-9	Sequence 9, App1	320	5	2.3	43	10	US-09-148-545-205	Sequence 205, App
248	5	2.3	21	10	US-09-853-830-40	Sequence 40, App1	321	5	2.3	43	10	US-09-147-761-3	Sequence 3, App1
249	5	2.3	21	10	US-09-853-830-50	Sequence 50, App1	322	5	2.3	43	10	US-09-843-845-21	Sequence 21, App1
250	5	2.3	21	10	US-09-853-830-64	Sequence 64, App1	323	5	2.3	43	10	US-09-864-761-37268	Sequence 37268, A
251	5	2.3	21	10	US-09-853-830-105	Sequence 105, App	324	5	2.3	43	10	US-09-864-761-43844	Sequence 43844, A
252	5	2.3	21	10	US-09-853-830-106	Sequence 106, App	325	5	2.3	43	10	US-09-925-300-1031	Sequence 1031, Ap
253	5	2.3	22	10	US-09-864-761-44266	Sequence 44266, A	326	5	2.3	44	10	US-09-843-845-6	Sequence 6, App1
254	5	2.3	22	10	US-09-326-248B-14	Sequence 14, App1	327	5	2.3	44	10	US-09-843-845-16	Sequence 16, App1
255	5	2.3	22	10	US-09-326-248B-17	Sequence 17, App1	328	5	2.3	44	10	US-09-864-761-34453	Sequence 34453, A
256	5	2.3	23	9	US-10-112-455A-10	Sequence 10, App1	329	5	2.3	44	10	US-09-864-761-36783	Sequence 36783, A
257	5	2.3	24	9	US-10-174-654-6	Sequence 6, App1	330	5	2.3	44	10	US-09-864-761-46624	Sequence 46624, A
258	5	2.3	24	9	US-09-991-548-42	Sequence 42, App1	331	5	2.3	45	9	US-09-925-299-1120	Sequence 1120, Ap
259	5	2.3	25	9	US-09-991-548-43	Sequence 43, App1	332	5	2.3	45	10	US-09-864-761-35639	Sequence 35639, A
260	5	2.3	25	9	US-09-991-548-44	Sequence 44, App1	333	5	2.3	45	10	US-09-864-761-36504	Sequence 36504, A
261	5	2.3	25	9	US-10-097-065-294	Sequence 294, App	334	5	2.3	45	10	US-09-864-761-39258	Sequence 39258, A
262	5	2.3	25	10	US-09-810-310-45	Sequence 45, App1	335	5	2.3	45	10	US-09-864-761-41747	Sequence 41747, A
263	5	2.3	25	10	US-09-864-761-40046	Sequence 40046, A	336	5	2.3	45	10	US-09-864-761-42078	Sequence 42078, A
264	5	2.3	25	10	US-09-402-131-8	Sequence 8, App1	337	5	2.3	45	10	US-09-925-299-1120	Sequence 4, App1
265	5	2.3	26	9	US-10-050-704-343	Sequence 343, App	338	5	2.3	46	10	US-09-864-761-37334	Sequence 37334, A
266	5	2.3	26	10	US-09-864-761-35286	Sequence 35286, A	339	5	2.3	46	10	US-09-864-761-41613	Sequence 41613, A
267	5	2.3	27	9	US-09-843-676-157	Sequence 157, App	340	5	2.3	46	10	US-09-864-761-41613	Sequence 359, App
268	5	2.3	27	9	US-09-843-676-169	Sequence 169, App	341	5	2.3	47	9	US-09-809-391-359	Sequence 354, App
269	5	2.3	27	9	US-09-766-253-157	Sequence 157, App	342	5	2.3	47	9	US-09-809-391-364	Sequence 34279, A
270	5	2.3	27	9	US-09-766-253-169	Sequence 169, App	343	5	2.3	47	10	US-09-864-761-33479	Sequence 34661, A
271	5	2.3	27	9	US-09-438-486-157	Sequence 157, App	344	5	2.3	47	10	US-09-864-761-34661	Sequence 48973, A
272	5	2.3	27	9	US-09-438-486-169	Sequence 169, App	345	5	2.3	47	10	US-09-864-761-48973	Sequence 288, App
273	5	2.3	27	9	US-10-053-758-157	Sequence 157, App	346	5	2.3	47	10	US-09-864-761-47351	Sequence 47351, A
274	5	2.3	27	9	US-10-053-758-169	Sequence 169, App	347	5	2.3	48	10	US-09-864-761-47351	Sequence 44784, A
275	5	2.3	27	9	US-10-054-295-157	Sequence 157, App	348	5	2.3	49	10	US-09-864-761-44784	Sequence 44033, A
276	5	2.3	27	9	US-10-054-295-169	Sequence 169, App	349	5	2.3	50	10	US-09-864-761-44033	Sequence 46791, A
277	5	2.3	27	9	US-10-054-611-157	Sequence 157, App	350	5	2.3	50	10	US-09-864-761-46791	Sequence 167, App
278	5	2.3	27	9	US-10-054-611-169	Sequence 169, App	351	5	2.3	50	10	US-09-071-838-167	Sequence 39450, A
279	5	2.3	28	9	US-09-974-879-573	Sequence 573, App	352	5	2.3	51	10	US-09-864-761-39450	Sequence 40963, A
280	5	2.3	29	10	US-09-864-761-45206	Sequence 45206, A	353	5	2.3	51	10	US-09-864-761-40963	Sequence 38811, A
281	5	2.3	29	10	US-09-864-761-47299	Sequence 47299, A	354	5	2.3	52	10	US-09-864-761-38811	Sequence 6708, Ap
282	5	2.3	30	9	US-09-840-277-53	Sequence 53, App1	355	5	2.3	53	9	US-09-738-626-6708	Sequence 44365, A
283	5	2.3	30	9	US-09-840-277-94	Sequence 94, App1	356	5	2.3	53	10	US-09-729-835-44	Sequence 333, App
284	5	2.3	30	10	US-09-864-761-41542	Sequence 41542, A	357	5	2.3	53	10	US-09-984-245-231	Sequence 1354, Ap
285	5	2.3	32	9	US-09-843-676-136	Sequence 136, App	358	5	2.3	54	9	US-09-984-245-231	Sequence 221, App
286	5	2.3	32	9	US-09-989-920-252	Sequence 252, App	359	5	2.3	54	9	US-09-966-262-231	Sequence 221, App
287	5	2.3	32	9	US-09-766-253-136	Sequence 136, App	360	5	2.3	54	9	US-09-983-966-231	Sequence 39008, A
288	5	2.3	32	9	US-09-438-486-136	Sequence 136, App	361	5	2.3	54	10	US-09-864-761-39008	Sequence 48954, A
289	5	2.3	32	9	US-10-053-758-136	Sequence 136, App	362	5	2.3	54	10	US-09-864-761-48954	Sequence 76, App1
290	5	2.3	32	9	US-10-054-295-136	Sequence 136, App	363	5	2.3	56	9	US-09-948-820-73	Sequence 209, App
291	5	2.3	32	9	US-10-054-611-136	Sequence 136, App	364	5	2.3	56	10	US-09-779-451-76	Sequence 42360, A
292	5	2.3	32	10	US-09-420-920-117	Sequence 117, App	365	5	2.3	57	10	US-09-879-957-209	Sequence 780, App
293	5	2.3	32	10	US-09-864-761-38821	Sequence 38821, A	366	5	2.3	57	10	US-09-864-761-44260	Sequence 33415, A
294	5	2.3	34	10	US-09-864-761-47407	Sequence 47407, A	367	5	2.3	58	10	US-09-867-550-780	Sequence 40365, A
295	5	2.3	34	10	US-09-864-761-48236	Sequence 48236, A	368	5	2.3	58	10	US-09-864-761-33415	Sequence 333, App
296	5	2.3	36	9	US-09-864-761-128	Sequence 128, App	369	5	2.3	58	10	US-09-864-761-40365	Sequence 1316, Ap
297	5	2.3	36	9	US-09-809-391-400	Sequence 400, App	370	5	2.3	59	9	US-10-108-605-353	Sequence 43, App1
298	5	2.3	36	10	US-09-864-761-36671	Sequence 36671, A	371	5	2.3	59	9	US-09-796-692-1354	Sequence 2320, Ap
299	5	2.3	36	10	US-09-864-761-39941	Sequence 39941, A	372	5	2.3	59	9	US-09-796-692-2320	Sequence 37468, A
300	5	2.3	36	10	US-09-864-761-42356	Sequence 42356, A	373	5	2.3	59	10	US-09-864-761-37468	Sequence 38914, A
301	5	2.3	37	9	US-10-042-141-184	Sequence 141, App	374	5	2.3	59	10	US-09-864-761-38914	Sequence 1316, Ap
302	5	2.3	37	9	US-10-050-704-194	Sequence 194, App	375	5	2.3	59	10	US-09-925-300-1316	Sequence 43, App1
303	5	2.3	37	10	US-09-726-643-184	Sequence 184, App	376	5	2.3	60	9	US-09-764-504-43	Sequence 35351, A
304	5	2.3	37	10	US-09-864-761-46637	Sequence 46637, A	377	5	2.3	60	9	US-10-091-548-43	Sequence 45360, A
305	5	2.3	38	10	US-09-864-761-33348	Sequence 33348, A	378	5	2.3	60	10	US-09-864-761-33351	Sequence 34950, A
306	5	2.3	38	10	US-09-864-761-40594	Sequence 40594, A	379	5	2.3	60	10	US-09-864-761-45360	Sequence 40688, A
307	5	2.3	38	10	US-09-864-761-47842	Sequence 47842, A	380	5	2.3	61	10	US-09-864-761-34950	Sequence 136, App
308	5	2.3	39	9	US-09-974-879-409	Sequence 937, App	381	5	2.3	61	10	US-09-864-761-34950	Sequence 136, App
309	5	2.3	39	9	US-10-092-154-937	Sequence 937, App	382	5	2.3	61	10	US-09-864-761-34950	Sequence 136, App
310	5	2.3	39	9	US-10-007-280A-208	Sequence 208, App	383	5	2.3	61	10	US-09-864-761-40688	Sequence 136, App
311	5	2.3	39	10	US-09-864-761-36173	Sequence 36173, A	384	5	2.3	62	9	US-09-996-140-136	Sequence 136, App



385	5	2.3	62	9	US-10-083-357-840	Sequence 840, App	458	5	2.3	86	9	US-10-072-783-3	Sequence 3, App1
386	5	2.3	62	10	US-09-764-877-1735	Sequence 1735, App	459	5	2.3	86	9	US-10-072-783-4	Sequence 4, App1
387	5	2.3	62	12	US-10-090-696-2	Sequence 2, App1	460	5	2.3	86	9	US-09-938-806A-2	Sequence 2, App1
388	5	2.3	63	9	US-09-984-245-232	Sequence 232, App	461	5	2.3	86	9	US-09-938-806A-3	Sequence 3, App1
389	5	2.3	63	9	US-09-966-262-232	Sequence 232, App	462	5	2.3	87	10	US-09-867-550-1648	Sequence 1648, App
390	5	2.3	63	9	US-09-983-966-232	Sequence 232, App	463	5	2.3	87	10	US-10-127-161-17	Sequence 17, App1
391	5	2.3	63	10	US-09-864-761-48562	Sequence 48562, App	464	5	2.3	89	9	US-10-027-629-17	Sequence 17, App1
392	5	2.3	63	10	US-09-925-301-1159	Sequence 1159, App	465	5	2.3	89	9	US-09-834-794A-34	Sequence 34, App1
393	5	2.3	64	10	US-09-925-301-1410	Sequence 1410, App	466	5	2.3	89	10	US-09-864-761-34424	Sequence 34424, App
394	5	2.3	65	10	US-09-864-761-42653	Sequence 42653, App	467	5	2.3	89	10	US-09-864-761-35769	Sequence 35769, App
395	5	2.3	65	10	US-09-925-297-460	Sequence 460, App	468	5	2.3	89	10	US-09-334-923A-6	Sequence 6, App1
396	5	2.3	65	10	US-09-956-194-16	Sequence 16, App1	469	5	2.3	89	10	US-09-334-954A-6	Sequence 6, App1
397	5	2.3	66	10	US-09-738-626-5959	Sequence 5959, App	470	5	2.3	89	10	US-09-334-954A-6	Sequence 6, App1
398	5	2.3	66	10	US-09-864-761-35040	Sequence 35040, App	471	5	2.3	90	10	US-09-334-954A-6	Sequence 6, App1
399	5	2.3	66	10	US-09-942-446-11	Sequence 11, App1	472	5	2.3	90	10	US-09-334-954A-6	Sequence 6, App1
400	5	2.3	66	10	US-09-764-877-1106	Sequence 1106, App	473	5	2.3	90	10	US-09-858-664A-12	Sequence 12, App1
401	5	2.3	67	8	US-08-808-031A-14	Sequence 14, App1	474	5	2.3	90	10	US-09-839-980-262	Sequence 262, App
402	5	2.3	67	10	US-09-864-761-36413	Sequence 36413, App	475	5	2.3	91	9	US-10-002-344A-195	Sequence 195, App
403	5	2.3	68	9	US-10-091-483-159	Sequence 159, App	476	5	2.3	91	9	US-10-042-431-32	Sequence 32, App1
404	5	2.3	68	10	US-09-764-846-159	Sequence 159, App	477	5	2.3	91	9	US-09-738-626-4813	Sequence 4813, App
405	5	2.3	69	9	US-10-001-835-167	Sequence 167, App	478	5	2.3	91	9	US-09-983-602-162	Sequence 162, App
406	5	2.3	69	9	US-09-792-793A-28	Sequence 28, App1	479	5	2.3	91	9	US-10-091-504-1239	Sequence 1239, App
407	5	2.3	70	9	US-10-011-445-73	Sequence 73, App1	480	5	2.3	91	10	US-09-864-761-43047	Sequence 43047, App
408	5	2.3	70	10	US-09-739-907-196	Sequence 196, App	481	5	2.3	91	10	US-09-764-869-1239	Sequence 1239, App
409	5	2.3	70	10	US-09-864-761-45889	Sequence 45889, App	482	5	2.3	93	8	US-08-424-508B-19	Sequence 19, App1
410	5	2.3	71	9	US-09-796-692-1501	Sequence 1501, App	483	5	2.3	93	8	US-10-102-806-616	Sequence 616, App
411	5	2.3	71	9	US-09-796-692-1520	Sequence 1520, App	484	5	2.3	93	10	US-09-864-761-36122	Sequence 36122, App
412	5	2.3	71	10	US-09-739-907-100	Sequence 100, App	485	5	2.3	93	10	US-09-867-550-590	Sequence 590, App
413	5	2.3	71	10	US-09-144-838-3	Sequence 3, App1	486	5	2.3	93	10	US-09-939-980-340	Sequence 340, App
414	5	2.3	72	10	US-09-864-761-34962	Sequence 34962, App	487	5	2.3	94	9	US-10-111-634-4	Sequence 4, App1
415	5	2.3	72	10	US-09-864-761-41955	Sequence 41955, App	488	5	2.3	94	10	US-09-739-907-2	Sequence 2, App1
416	5	2.3	73	9	US-10-178-213-5	Sequence 5, App1	489	5	2.3	95	9	US-09-866-050A-488	Sequence 488, App
417	5	2.3	73	9	US-09-809-391-579	Sequence 8, App1	490	5	2.3	95	10	US-09-864-761-42885	Sequence 42885, App
418	5	2.3	73	10	US-09-864-761-37231	Sequence 37231, App	491	5	2.3	95	10	US-09-764-868-871	Sequence 871, App
419	5	2.3	73	10	US-09-144-838-6	Sequence 6, App1	492	5	2.3	96	9	US-09-911-777-4	Sequence 4, App1
420	5	2.3	73	10	US-09-867-550-306	Sequence 306, App	493	5	2.3	96	10	US-09-764-877-1656	Sequence 1656, App
421	5	2.3	73	10	US-09-764-877-1680	Sequence 1680, App	494	5	2.3	96	10	US-09-263-959-1201	Sequence 1201, App
422	5	2.3	74	9	US-10-102-806-800	Sequence 800, App	495	5	2.3	97	9	US-09-776-724A-210	Sequence 210, App
423	5	2.3	74	10	US-09-864-761-47484	Sequence 47484, App	496	5	2.3	97	10	US-09-925-902-792	Sequence 792, App
424	5	2.3	74	10	US-09-764-877-1504	Sequence 1504, App	497	5	2.3	97	10	US-09-764-877-1526	Sequence 1526, App
425	5	2.3	74	10	US-09-925-300-1586	Sequence 1586, App	498	5	2.3	98	10	US-09-846-258-4	Sequence 4, App1
426	5	2.3	75	9	US-10-092-154-834	Sequence 834, App	499	5	2.3	98	10	US-09-864-761-35850	Sequence 35850, App
427	5	2.3	75	10	US-09-764-847-834	Sequence 834, App	500	5	2.3	98	10	US-09-893-238-126	Sequence 126, App
428	5	2.3	76	9	US-09-955-999-88	Sequence 88, App1	501	5	2.3	100	10	US-09-867-550-1160	Sequence 1160, App
429	5	2.3	76	10	US-09-864-761-39803	Sequence 39803, App	502	5	2.3	101	10	US-09-738-626-4790	Sequence 4790, App
430	5	2.3	76	10	US-09-864-761-44389	Sequence 44389, App	503	5	2.3	102	10	US-09-864-761-36527	Sequence 36527, App
431	5	2.3	76	10	US-09-764-877-1436	Sequence 1436, App	504	5	2.3	102	10	US-09-216-393-135	Sequence 135, App
432	5	2.3	77	9	US-09-984-245-273	Sequence 273, App	505	5	2.3	103	10	US-09-815-542-5625	Sequence 5625, App
433	5	2.3	77	9	US-09-738-626-6969	Sequence 6969, App	506	5	2.3	103	10	US-09-815-542-12246	Sequence 12246, App
434	5	2.3	77	9	US-10-178-213-437	Sequence 437, App	507	5	2.3	103	10	US-09-815-542-12246	Sequence 12246, App
435	5	2.3	77	9	US-09-966-262-273	Sequence 273, App	508	5	2.3	103	10	US-09-815-542-12246	Sequence 12246, App
436	5	2.3	77	9	US-09-983-966-273	Sequence 273, App	509	5	2.3	104	9	US-09-834-794A-2	Sequence 2, App1
437	5	2.3	77	10	US-09-864-761-44275	Sequence 44275, App	510	5	2.3	104	9	US-10-116-016-28	Sequence 28, App1
438	5	2.3	78	10	US-10-016-157A-162	Sequence 162, App	511	5	2.3	104	10	US-09-834-795A-2	Sequence 2, App1
439	5	2.3	78	10	US-09-864-761-34713	Sequence 34713, App	512	5	2.3	104	10	US-09-764-848-28	Sequence 28, App1
440	5	2.3	78	10	US-09-939-980-503	Sequence 503, App	513	5	2.3	105	9	US-09-738-626-7321	Sequence 7321, App
441	5	2.3	78	10	US-10-158-366-6	Sequence 6, App1	514	5	2.3	106	9	US-10-102-806-732	Sequence 732, App
442	5	2.3	79	10	US-09-864-761-35205	Sequence 35205, App	515	5	2.3	106	10	US-09-739-907-195	Sequence 195, App
443	5	2.3	79	10	US-09-864-761-44436	Sequence 44436, App	516	5	2.3	106	10	US-09-846-258-5	Sequence 5, App1
444	5	2.3	80	10	US-09-864-761-44436	Sequence 44436, App	517	5	2.3	109	9	US-09-764-868-795	Sequence 795, App
445	5	2.3	81	9	US-09-738-626-6898	Sequence 6898, App	518	5	2.3	109	9	US-09-989-442-143	Sequence 143, App
446	5	2.3	81	9	US-09-974-879-571	Sequence 571, App	519	5	2.3	109	9	US-09-969-730-165	Sequence 165, App
447	5	2.3	81	9	US-09-866-050A-490	Sequence 490, App	520	5	2.3	109	9	US-10-091-504-1060	Sequence 1060, App
448	5	2.3	81	10	US-09-729-674-80	Sequence 80, App1	521	5	2.3	109	10	US-09-764-869-1060	Sequence 1060, App
449	5	2.3	81	10	US-09-864-761-37861	Sequence 37861, App	522	5	2.3	110	9	US-09-774-639-112	Sequence 112, App
450	5	2.3	81	10	US-09-864-761-44504	Sequence 44504, App	523	5	2.3	110	10	US-09-864-761-47631	Sequence 47631, App
451	5	2.3	82	10	US-10-102-806-780	Sequence 780, App	524	5	2.3	111	9	US-09-974-879-559	Sequence 559, App
452	5	2.3	82	10	US-09-731-872-273	Sequence 273, App	525	5	2.3	111	10	US-09-263-959-349	Sequence 349, App
453	5	2.3	83	9	US-09-738-626-6475	Sequence 6475, App	526	5	2.3	113	9	US-09-984-245-233	Sequence 233, App
454	5	2.3	84	9	US-09-843-221A-11	Sequence 11, App1	527	5	2.3	113	9	US-09-764-868-868	Sequence 868, App
455	5	2.3	84	10	US-09-864-761-46572	Sequence 46572, App	528	5	2.3	113	9	US-09-764-868-1179	Sequence 1179, App
456	5	2.3	84	10	US-09-867-550-1206	Sequence 1206, App	529	5	2.3	113	9	US-09-966-262-233	Sequence 233, App
457	5	2.3	85	9	US-10-083-357-1259	Sequence 1259, App	530	5	2.3	113	9	US-10-092-154-678	Sequence 678, App

531	5	2.3	113	9	US-09-993-966-233	Sequence 233, App	604	5	2.3	142	9	US-10-102-806-799	Sequence 799, App
532	5	2.3	113	10	US-09-764-847-678	Sequence 678, App	605	5	2.3	143	9	US-09-764-668-666	Sequence 666, App
533	5	2.3	114	9	US-09-866-050A-161	Sequence 161, App	606	5	2.3	143	9	US-09-738-626-5363	Sequence 5363, App
534	5	2.3	114	9	US-09-866-050A-288	Sequence 288, App	607	5	2.3	143	9	US-09-738-626-5960	Sequence 5960, App
535	5	2.3	114	10	US-09-864-761-45709	Sequence 45709, A	608	5	2.3	143	9	US-09-796-692-1372	Sequence 1372, App
536	5	2.3	115	9	US-09-738-626-6211	Sequence 6211, App	609	5	2.3	143	9	US-09-796-692-2339	Sequence 2339, App
537	5	2.3	115	9	US-10-091-572-429	Sequence 229, App	610	5	2.3	143	10	US-09-731-872-2772	Sequence 2772, App
538	5	2.3	115	9	US-10-091-572-429	Sequence 1052, App	611	5	2.3	143	10	US-09-925-300-1867	Sequence 1867, App
539	5	2.3	115	10	US-09-764-869-1052	Sequence 1052, App	612	5	2.3	145	9	US-09-949-442-16	Sequence 16, App1
540	5	2.3	116	9	US-10-043-487-240	Sequence 240, App	613	5	2.3	145	10	US-09-894-526-1	Sequence 1, App1
541	5	2.3	116	10	US-09-854-845-35	Sequence 35, App1	614	5	2.3	146	9	US-10-172-088-10	Sequence 10, App1
542	5	2.3	117	9	US-09-789-390-73	Sequence 73, App1	615	5	2.3	146	10	US-09-925-301-1214	Sequence 1214, App
543	5	2.3	117	12	US-10-022-461-6	Sequence 6, App1	616	5	2.3	146	10	US-09-764-864-1147	Sequence 1147, App
544	5	2.3	118	9	US-09-864-761-41420	Sequence 41420, A	617	5	2.3	147	10	US-09-819-994-2	Sequence 2, App1
545	5	2.3	119	9	US-10-163-316-4	Sequence 4, App1	618	5	2.3	147	10	US-09-864-761-36066	Sequence 36066, A
546	5	2.3	119	10	US-09-216-393-2	Sequence 2, App1	619	5	2.3	148	9	US-09-796-692-1364	Sequence 1364, App
547	5	2.3	119	10	US-09-305-658-246	Sequence 246, App	620	5	2.3	148	9	US-10-013-379-9	Sequence 9, App1
548	5	2.3	119	10	US-09-858-081-4	Sequence 4, App1	621	5	2.3	149	9	US-09-738-626-4907	Sequence 4907, App
549	5	2.3	119	10	US-09-858-068-4	Sequence 4, App1	622	5	2.3	149	10	US-09-815-242-11049	Sequence 11049, A
550	5	2.3	119	12	US-10-014-070-7	Sequence 7, App1	623	5	2.3	149	10	US-09-764-853-839	Sequence 839, App
551	5	2.3	120	10	US-09-864-761-46706	Sequence 46706, A	624	5	2.3	150	9	US-10-042-431-29	Sequence 29, App1
552	5	2.3	120	10	US-09-764-877-1108	Sequence 1108, App	625	5	2.3	150	9	US-10-042-431-44	Sequence 44, App1
553	5	2.3	121	9	US-09-738-626-6781	Sequence 6781, App	626	5	2.3	150	9	US-09-759-1308A-414	Sequence 414, App
554	5	2.3	122	9	US-10-042-141-91	Sequence 9, App1	627	5	2.3	151	10	US-09-746-359A-3	Sequence 3, App1
555	5	2.3	122	9	US-09-983-802-285	Sequence 285, App	628	5	2.3	151	10	US-09-515-806-29	Sequence 29, App1
556	5	2.3	122	10	US-09-726-643-91	Sequence 91, App1	629	5	2.3	152	10	US-09-741-669-445	Sequence 445, App
557	5	2.3	122	10	US-09-908-711-119	Sequence 119, App	630	5	2.3	152	10	US-09-746-359A-2	Sequence 2, App1
558	5	2.3	122	10	US-09-815-242-11436	Sequence 11436, A	631	5	2.3	153	10	US-09-815-242-10946	Sequence 10946, A
559	5	2.3	122	10	US-09-815-242-12001	Sequence 12001, A	632	5	2.3	153	10	US-09-825-300-1688	Sequence 1688, App
560	5	2.3	123	9	US-09-796-652-1360	Sequence 1360, App	633	5	2.3	154	10	US-09-925-301-851	Sequence 851, App
561	5	2.3	123	9	US-09-796-632-2329	Sequence 2329, App	634	5	2.3	155	10	US-09-593-634-21	Sequence 21, App1
562	5	2.3	123	9	US-10-043-487-333	Sequence 333, App	635	5	2.3	155	10	US-09-925-297-504	Sequence 504, App
563	5	2.3	124	9	US-10-101-464A-567	Sequence 567, App	636	5	2.3	155	10	US-09-917-340-29	Sequence 29, App1
564	5	2.3	125	9	US-09-796-692-1301	Sequence 1301, App	637	5	2.3	156	9	US-09-854-133-120	Sequence 20, App
565	5	2.3	125	9	US-09-796-692-2292	Sequence 2292, App	638	5	2.3	156	9	US-09-796-692-2357	Sequence 2357, App
566	5	2.3	125	10	US-09-864-761-39406	Sequence 39406, A	639	5	2.3	156	9	US-09-977-418-38	Sequence 38, App1
567	5	2.3	126	10	US-09-925-302-699	Sequence 699, App	640	5	2.3	156	10	US-09-823-356-5	Sequence 5, App1
568	5	2.3	127	9	US-09-898-751A-6	Sequence 6, App1	641	5	2.3	156	10	US-09-738-973-210	Sequence 210, App
569	5	2.3	127	9	US-09-834-794A-1	Sequence 1, App1	642	5	2.3	156	10	US-09-771-161A-104	Sequence 104, App
570	5	2.3	127	9	US-10-146-496-2	Sequence 2, App1	643	5	2.3	156	12	US-10-027-450-6	Sequence 6, App1
571	5	2.3	127	10	US-09-813-492-2	Sequence 2, App1	644	5	2.3	157	9	US-09-764-868-648	Sequence 648, App
572	5	2.3	127	10	US-09-746-359A-4	Sequence 4, App1	645	5	2.3	157	9	US-09-764-868-649	Sequence 649, App
573	5	2.3	127	10	US-09-811-284-222	Sequence 222, App	646	5	2.3	157	10	US-09-811-284-204	Sequence 204, App
574	5	2.3	127	10	US-09-834-795A-1	Sequence 1, App1	647	5	2.3	158	9	US-09-738-626-5695	Sequence 5695, App
575	5	2.3	127	10	US-09-867-550-1666	Sequence 1666, App	648	5	2.3	158	9	US-09-796-692-1356	Sequence 1356, App
576	5	2.3	127	10	US-09-931-381A-2	Sequence 2, App1	649	5	2.3	158	9	US-09-796-692-2325	Sequence 2325, App
577	5	2.3	130	9	US-09-898-751A-10	Sequence 10, App1	650	5	2.3	159	9	US-09-809-391-427	Sequence 427, App
578	5	2.3	130	9	US-10-060-845-11	Sequence 11, App1	651	5	2.3	159	9	US-10-116-016-33	Sequence 33, App1
579	5	2.3	130	9	US-10-060-845-12	Sequence 12, App1	652	5	2.3	159	10	US-09-764-848-33	Sequence 33, App1
580	5	2.3	130	9	US-10-060-845-13	Sequence 13, App1	653	5	2.3	159	12	US-10-062-254-342	Sequence 342, App
581	5	2.3	130	9	US-10-102-806-685	Sequence 685, App	654	5	2.3	160	9	US-09-738-626-5386	Sequence 5386, App
582	5	2.3	130	12	US-10-072-159-11	Sequence 11, App1	655	5	2.3	160	10	US-09-747-155-154	Sequence 154, App
583	5	2.3	130	12	US-10-072-159-12	Sequence 12, App1	656	5	2.3	160	10	US-09-747-155-154	Sequence 154, App
584	5	2.3	130	12	US-10-072-159-13	Sequence 13, App1	657	5	2.3	161	9	US-10-092-154-593	Sequence 593, App
585	5	2.3	131	10	US-09-815-242-10458	Sequence 10458, A	658	5	2.3	161	10	US-09-745-003-9	Sequence 9, App1
586	5	2.3	131	10	US-09-815-242-13858	Sequence 13858, A	659	5	2.3	161	10	US-09-745-003-9	Sequence 9, App1
587	5	2.3	131	10	US-09-825-414-58	Sequence 58, App1	660	5	2.3	161	10	US-09-764-847-593	Sequence 593, App
588	5	2.3	131	10	US-09-781-811-2	Sequence 2, App1	661	5	2.3	162	9	US-09-738-626-4267	Sequence 4267, App
589	5	2.3	132	10	US-09-840-787-38	Sequence 38, App1	662	5	2.3	162	10	US-09-861-451A-163	Sequence 4667, App1
590	5	2.3	132	10	US-09-866-582-23	Sequence 23, App1	663	5	2.3	162	10	US-09-925-300-1362	Sequence 1362, App
591	5	2.3	132	10	US-09-781-811-4	Sequence 4, App1	664	5	2.3	162	10	US-09-747-155-194	Sequence 194, App
592	5	2.3	133	9	US-09-738-626-6471	Sequence 6471, App	665	5	2.3	163	9	US-09-738-626-1129	Sequence 4129, App
593	5	2.3	133	9	US-09-738-626-6718	Sequence 6718, App	666	5	2.3	163	9	US-09-861-451A-80	Sequence 80, App1
594	5	2.3	135	10	US-09-929-493-20	Sequence 20, App1	667	5	2.3	163	10	US-09-764-864-962	Sequence 962, App
595	5	2.3	137	10	US-09-740-046-6	Sequence 6, App1	668	5	2.3	163	10	US-09-764-864-1563	Sequence 1563, App
596	5	2.3	137	10	US-09-867-550-524	Sequence 524, App	669	5	2.3	165	9	US-10-043-487-122	Sequence 322, App
597	5	2.3	139	10	US-09-819-994-4	Sequence 4, App1	670	5	2.3	166	9	US-10-097-065-242	Sequence 95, App1
598	5	2.3	139	10	US-09-827-949-31	Sequence 31, App1	671	5	2.3	166	10	US-09-815-242-11642	Sequence 11642, A
599	5	2.3	140	10	US-09-864-641-2	Sequence 2, App1	672	5	2.3	167	10	US-09-739-254-125	Sequence 125, App
600	5	2.3	141	9	US-09-796-692-1310	Sequence 1783, App	673	5	2.3	167	10	US-09-904-615-125	Sequence 125, App
601	5	2.3	141	9	US-09-796-692-1783	Sequence 2051, App	674	5	2.3	168	9	US-09-981-676-880	Sequence 280, App
602	5	2.3	141	9	US-09-796-692-2051	Sequence 2301, App	675	5	2.3	168	9	US-09-148-545-280	Sequence 280, App
603	5	2.3	141	9	US-09-796-692-2301		676	5	2.3	168	9		

677	5	2.3	168	10	US-09-879-919-4	Sequence 4, Appli	750	5	2.3	179	9	US-10-066-273-126	Sequence 126, App
678	5	2.3	168	12	US-09-731-872-324	Sequence 324, App	751	5	2.3	179	9	US-10-066-494-126	Sequence 126, App
679	5	2.3	168	10	US-10-082-260-4	Sequence 4, Appli	752	5	2.3	179	9	US-10-230-163-244	Sequence 244, App
680	5	2.3	168	12	US-10-095-492-25	Sequence 25, Appl	753	5	2.3	179	9	US-10-066-269-126	Sequence 126, App
681	5	2.3	169	6	US-09-738-626-5504	Sequence 5504, Ap	754	5	2.3	179	9	US-10-066-193-126	Sequence 126, App
682	5	2.3	169	9	US-09-729-674-56	Sequence 56, Appl	755	5	2.3	179	9	US-10-066-211-126	Sequence 126, App
683	5	2.3	169	10	US-09-864-761-48249	Sequence 48249, A	756	5	2.3	179	9	US-10-218-631-244	Sequence 244, App
684	5	2.3	170	9	US-10-078-770-8	Sequence 8, Appli	757	5	2.3	179	9	US-10-230-338-444	Sequence 444, App
685	5	2.3	170	9	US-10-026-741-65	Sequence 65, Appl	758	5	2.3	179	9	US-10-063-518-154	Sequence 154, App
686	5	2.3	171	9	US-10-026-741-96	Sequence 96, Appl	759	5	2.3	179	9	US-10-063-598-154	Sequence 154, App
687	5	2.3	171	10	US-09-771-161A-124	Sequence 124, App	760	5	2.3	179	9	US-10-227-693-154	Sequence 154, App
688	5	2.3	172	9	US-09-779-050A-17	Sequence 17, Appl	761	5	2.3	179	9	US-10-230-414-244	Sequence 244, App
689	5	2.3	172	9	US-10-042-431-27	Sequence 27, Appl	762	5	2.3	179	10	US-09-799-160-4	Sequence 4, Appli
690	5	2.3	172	9	US-10-042-431-74	Sequence 74, Appl	763	5	2.3	179	10	US-09-751-797-27	Sequence 27, Appl
691	5	2.3	172	10	US-09-816-494-8	Sequence 8, Appli	764	5	2.3	179	10	US-09-751-797-28	Sequence 28, Appl
692	5	2.3	172	10	US-09-925-301-894	Sequence 894, App	765	5	2.3	179	10	US-09-728-911-15	Sequence 15, Appl
693	5	2.3	172	10	US-09-815-419-5	Sequence 5, Appli	766	5	2.3	179	10	US-09-870-574-2	Sequence 2, Appli
694	5	2.3	173	9	US-10-135-807-3	Sequence 3, Appli	767	5	2.3	179	12	US-10-006-867-154	Sequence 154, App
695	5	2.3	173	10	US-09-816-494-7	Sequence 7, Appli	768	5	2.3	180	9	US-09-796-682-2361	Sequence 2361, Ap
696	5	2.3	173	10	US-09-815-419-4	Sequence 4, Appli	769	5	2.3	180	9	US-10-131-241-17	Sequence 17, Appl
697	5	2.3	173	10	US-09-939-980-476	Sequence 476, App	770	5	2.3	180	9	US-10-131-241-16	Sequence 16, Appl
698	5	2.3	174	9	US-09-809-391-424	Sequence 424, App	771	5	2.3	180	10	US-09-939-980-337	Sequence 337, App
699	5	2.3	175	9	US-09-738-626-3664	Sequence 3664, Ap	772	5	2.3	181	9	US-09-738-626-5074	Sequence 5074, App
700	5	2.3	175	9	US-09-738-626-3818	Sequence 3818, Ap	773	5	2.3	181	9	US-10-131-241-55	Sequence 55, Appl
701	5	2.3	175	9	US-10-101-464A-653	Sequence 653, App	774	5	2.3	181	10	US-09-815-242-10433	Sequence 10433, A
702	5	2.3	176	9	US-09-984-245-191	Sequence 191, App	775	5	2.3	181	10	US-09-731-872-295	Sequence 295, Appl
703	5	2.3	176	9	US-10-028-072-98	Sequence 98, Appl	776	5	2.3	181	10	US-09-927-112-14	Sequence 14, Appl
704	5	2.3	176	9	US-10-121-049-98	Sequence 98, Appl	777	5	2.3	182	9	US-10-131-241-54	Sequence 54, Appl
705	5	2.3	176	9	US-10-123-904-98	Sequence 98, Appl	778	5	2.3	182	10	US-09-815-242-5254	Sequence 5254, Ap
706	5	2.3	176	9	US-10-140-470-98	Sequence 98, Appl	779	5	2.3	182	10	US-09-815-242-12452	Sequence 12452, A
707	5	2.3	176	9	US-10-175-746-98	Sequence 98, Appl	780	5	2.3	182	10	US-09-815-242-12786	Sequence 12786, A
708	5	2.3	176	9	US-10-176-921-98	Sequence 98, Appl	781	5	2.3	182	10	US-09-815-242-13054	Sequence 13054, A
709	5	2.3	176	9	US-10-176-921-98	Sequence 98, Appl	782	5	2.3	182	10	US-09-815-242-13054	Sequence 13054, A
710	5	2.3	176	9	US-10-227-884-218	Sequence 218, App	783	5	2.3	183	9	US-10-080-797-1	Sequence 1, Appli
711	5	2.3	176	9	US-10-137-865-98	Sequence 98, Appl	784	5	2.3	183	9	US-10-131-241-52	Sequence 52, Appl
712	5	2.3	176	9	US-10-140-471-98	Sequence 98, Appl	785	5	2.3	183	9	US-10-084-205-10	Sequence 10, Appl
713	5	2.3	176	9	US-10-142-431-98	Sequence 98, Appl	786	5	2.3	183	9	US-10-102-806-530	Sequence 530, App
714	5	2.3	176	9	US-10-143-114-98	Sequence 98, Appl	787	5	2.3	183	10	US-09-815-242-11840	Sequence 11840, A
715	5	2.3	176	9	US-10-230-163-218	Sequence 218, App	788	5	2.3	183	10	US-09-873-676-2	Sequence 2, Appli
716	5	2.3	176	9	US-10-142-431-98	Sequence 98, Appl	789	5	2.3	183	10	US-09-925-637-10	Sequence 10, Appl
717	5	2.3	176	9	US-10-142-431-98	Sequence 98, Appl	790	5	2.3	184	9	US-10-131-241-49	Sequence 49, Appl
718	5	2.3	176	9	US-10-218-631-218	Sequence 218, App	791	5	2.3	184	10	US-09-893-737-168	Sequence 168, App
719	5	2.3	176	9	US-10-230-338-218	Sequence 218, App	792	5	2.3	185	10	US-09-925-302-660	Sequence 660, App
720	5	2.3	176	9	US-09-966-262-191	Sequence 191, App	793	5	2.3	185	10	US-09-864-761-48667	Sequence 48667, A
721	5	2.3	176	9	US-10-123-262-98	Sequence 98, Appl	794	5	2.3	185	10	US-09-815-242-11083	Sequence 11083, A
722	5	2.3	176	9	US-10-142-423-98	Sequence 98, Appl	795	5	2.3	186	9	US-09-944-160-16	Sequence 16, Appl
723	5	2.3	176	9	US-10-230-414-218	Sequence 218, App	796	5	2.3	186	9	US-09-738-626-1639	Sequence 3639, Ap
724	5	2.3	176	9	US-10-121-050-98	Sequence 98, Appl	797	5	2.3	186	10	US-09-741-669-359	Sequence 359, App
725	5	2.3	176	9	US-10-141-755-98	Sequence 98, Appl	798	5	2.3	187	10	US-09-761-288-45	Sequence 45, Appl
726	5	2.3	176	9	US-09-983-966-191	Sequence 191, App	799	5	2.3	187	10	US-09-761-288-51	Sequence 51, Appl
727	5	2.3	176	9	US-10-143-032-98	Sequence 98, Appl	800	5	2.3	187	12	US-10-006-492-1	Sequence 1, Appli
728	5	2.3	176	10	US-09-746-359A-1	Sequence 1, Appli	801	5	2.3	190	9	US-10-078-770-26	Sequence 26, Appl
729	5	2.3	177	9	US-09-854-133-202	Sequence 202, App	802	5	2.3	190	9	US-10-050-704-342	Sequence 342, App
730	5	2.3	177	9	US-09-854-133-208	Sequence 208, App	803	5	2.3	191	9	US-09-789-380-74	Sequence 74, Appl
731	5	2.3	177	10	US-09-815-242-12000	Sequence 12000, A	804	5	2.3	191	10	US-09-867-550-1628	Sequence 1628, Ap
732	5	2.3	177	10	US-09-800-729-186	Sequence 186, App	805	5	2.3	193	9	US-09-796-753-162	Sequence 162, App
733	5	2.3	177	10	US-09-738-973-202	Sequence 202, App	806	5	2.3	193	10	US-09-864-761-34304	Sequence 34304, A
734	5	2.3	177	10	US-09-738-973-208	Sequence 208, App	807	5	2.3	193	10	US-09-864-761-37748	Sequence 37748, A
735	5	2.3	178	9	US-10-131-241-60	Sequence 60, Appl	808	5	2.3	193	10	US-09-841-132-581	Sequence 581, App
736	5	2.3	178	9	US-10-135-807-5	Sequence 5, Appli	809	5	2.3	193	10	US-09-764-853-673	Sequence 673, App
737	5	2.3	178	10	US-09-363-993-5	Sequence 5, Appli	810	5	2.3	194	9	US-10-108-605-23	Sequence 23, Appl
738	5	2.3	178	10	US-09-782-980-6	Sequence 6, Appli	811	5	2.3	194	9	US-10-227-884-42	Sequence 42, Appl
739	5	2.3	179	9	US-10-066-500-126	Sequence 126, App	812	5	2.3	194	9	US-10-230-163-32	Sequence 32, Appl
740	5	2.3	179	9	US-10-063-547-154	Sequence 154, App	813	5	2.3	194	9	US-10-218-631-32	Sequence 32, Appl
741	5	2.3	179	9	US-09-965-528-18	Sequence 18, Appl	814	5	2.3	194	9	US-10-230-338-12	Sequence 32, Appl
742	5	2.3	179	9	US-10-050-552A-2	Sequence 2, Appli	815	5	2.3	194	9	US-10-230-414-32	Sequence 32, Appl
743	5	2.3	179	9	US-09-738-626-4399	Sequence 4399, Ap	816	5	2.3	194	10	US-09-925-501-999	Sequence 999, App
744	5	2.3	179	9	US-09-738-626-4719	Sequence 4719, Ap	817	5	2.3	195	9	US-09-738-626-385	Sequence 4355, Ap
745	5	2.3	179	9	US-10-063-616-154	Sequence 154, App	818	5	2.3	196	9	US-10-025-367-4	Sequence 4, Appli
746	5	2.3	179	9	US-10-131-241-57	Sequence 57, Appl	819	5	2.3	196	9	US-10-028-072-168	Sequence 168, App
747	5	2.3	179	9	US-10-063-502-154	Sequence 154, App	820	5	2.3	196	9	US-10-121-049-168	Sequence 168, App
748	5	2.3	179	9	US-10-227-884-244	Sequence 244, App	821	5	2.3	196	9	US-10-123-904-168	Sequence 168, App
749	5	2.3	179	9	US-10-002-796-126	Sequence 126, App	822	5	2.3	196	9	US-10-140-470-168	Sequence 168, App

823	5	2.3	196	9	US-10-175-746-168	Sequence 168, App	896	5	2.3	211	9	US-10-174-590-102	Sequence 102, App
824	5	2.3	196	9	US-10-176-918-168	Sequence 168, App	897	5	2.3	211	9	US-10-176-758-102	Sequence 102, App
825	5	2.3	196	9	US-10-176-921-168	Sequence 168, App	898	5	2.3	211	9	US-10-175-737-102	Sequence 102, App
826	5	2.3	196	9	US-10-137-865-168	Sequence 168, App	899	5	2.3	211	9	US-10-173-706-102	Sequence 102, App
827	5	2.3	196	9	US-10-140-474-168	Sequence 168, App	900	5	2.3	211	9	US-10-175-738-102	Sequence 102, App
828	5	2.3	196	9	US-10-142-431-168	Sequence 168, App	901	5	2.3	211	9	US-10-175-752-102	Sequence 102, App
829	5	2.3	196	9	US-10-143-114-168	Sequence 168, App	902	5	2.3	211	9	US-10-176-482-102	Sequence 102, App
830	5	2.3	196	9	US-10-140-002-168	Sequence 168, App	903	5	2.3	211	9	US-10-176-757-102	Sequence 102, App
831	5	2.3	196	9	US-10-142-419-168	Sequence 168, App	904	5	2.3	211	9	US-10-176-913-102	Sequence 102, App
832	5	2.3	196	9	US-09-791-279-196	Sequence 196, App	905	5	2.3	211	9	US-10-180-552-102	Sequence 102, App
833	5	2.3	196	9	US-10-050-704-341	Sequence 341, App	906	5	2.3	211	9	US-10-180-557-102	Sequence 102, App
834	5	2.3	196	9	US-10-123-262-168	Sequence 168, App	907	5	2.3	211	9	US-10-173-700-102	Sequence 102, App
835	5	2.3	196	9	US-10-142-423-168	Sequence 168, App	908	5	2.3	211	9	US-10-174-572-102	Sequence 102, App
836	5	2.3	196	9	US-10-121-050-168	Sequence 168, App	909	5	2.3	211	9	US-10-174-579-102	Sequence 102, App
837	5	2.3	196	9	US-10-141-755-168	Sequence 168, App	910	5	2.3	211	9	US-10-174-582-102	Sequence 102, App
838	5	2.3	196	9	US-10-125-540-396	Sequence 396, App	911	5	2.3	211	9	US-10-174-588-102	Sequence 102, App
839	5	2.3	196	9	US-10-143-032-168	Sequence 168, App	912	5	2.3	211	9	US-10-175-739-102	Sequence 102, App
840	5	2.3	196	10	US-09-764-870-396	Sequence 396, App	913	5	2.3	211	9	US-10-175-740-102	Sequence 102, App
841	5	2.3	197	9	US-09-738-626-4953	Sequence 4953, App	914	5	2.3	211	9	US-10-175-743-102	Sequence 102, App
842	5	2.3	197	9	US-10-101-464A-491	Sequence 491, App	915	5	2.3	211	9	US-10-176-488-102	Sequence 102, App
843	5	2.3	197	10	US-09-881-752A-98	Sequence 98, Appl	916	5	2.3	211	9	US-10-176-492-102	Sequence 102, App
844	5	2.3	198	10	US-09-799-848-5	Sequence 5, Appl	917	5	2.3	211	9	US-10-176-747-102	Sequence 102, App
845	5	2.3	198	10	US-09-925-301-1184	Sequence 1184, Ap	918	5	2.3	211	9	US-10-176-750-102	Sequence 102, App
846	5	2.3	198	10	US-09-858-664A-9	Sequence 9, Appl	919	5	2.3	211	9	US-10-176-985-102	Sequence 102, App
847	5	2.3	199	9	US-09-738-626-4417	Sequence 4417, Ap	920	5	2.3	211	9	US-10-176-987-102	Sequence 102, App
848	5	2.3	199	9	US-09-738-626-4672	Sequence 4672, Ap	921	5	2.3	211	9	US-10-176-991-102	Sequence 102, App
849	5	2.3	199	9	US-10-125-540-489	Sequence 489, App	922	5	2.3	211	9	US-10-176-992-102	Sequence 102, App
850	5	2.3	199	10	US-09-764-870-489	Sequence 489, App	923	5	2.3	211	9	US-10-176-993-102	Sequence 102, App
851	5	2.3	199	10	US-09-953-260-2	Sequence 2, Appl	924	5	2.3	211	9	US-10-184-658-102	Sequence 102, App
852	5	2.3	199	10	US-09-764-853-850	Sequence 850, App	925	5	2.3	211	9	US-10-173-695-102	Sequence 102, App
853	5	2.3	199	10	US-09-731-909A-2	Sequence 2, Appl	926	5	2.3	211	9	US-10-173-697-102	Sequence 102, App
854	5	2.3	199	10	US-09-939-825-28	Sequence 28, Appl	927	5	2.3	211	9	US-10-173-705-102	Sequence 102, App
855	5	2.3	199	12	US-10-025-676-7	Sequence 7, Appl	928	5	2.3	211	9	US-10-174-576-102	Sequence 102, App
856	5	2.3	200	9	US-09-738-626-5845	Sequence 5845, Ap	929	5	2.3	211	9	US-10-174-585-102	Sequence 102, App
857	5	2.3	200	10	US-09-995-593A-2	Sequence 2, Appl	930	5	2.3	211	9	US-10-174-586-102	Sequence 102, App
858	5	2.3	201	9	US-10-226-408-1	Sequence 1, Appl	931	5	2.3	211	9	US-10-175-747-102	Sequence 102, App
859	5	2.3	201	9	US-10-177-916-1	Sequence 1, Appl	932	5	2.3	211	9	US-10-176-481-102	Sequence 102, App
860	5	2.3	201	9	US-09-746-660A-84	Sequence 84, Appl	933	5	2.3	211	9	US-10-176-485-102	Sequence 102, App
861	5	2.3	201	9	US-09-791-279-127	Sequence 127, App	934	5	2.3	211	9	US-10-176-487-102	Sequence 102, App
862	5	2.3	201	10	US-09-014-348-7	Sequence 7, Appl	935	5	2.3	211	9	US-10-176-493-102	Sequence 102, App
863	5	2.3	201	10	US-09-846-729A-1	Sequence 1, Appl	936	5	2.3	211	9	US-10-176-756-102	Sequence 102, App
864	5	2.3	201	10	US-09-782-980-37	Sequence 37, Appl	937	5	2.3	211	9	US-10-176-911-102	Sequence 102, App
865	5	2.3	202	10	US-09-846-729A-16	Sequence 16, Appl	938	5	2.3	211	9	US-10-176-919-102	Sequence 102, App
866	5	2.3	202	10	US-09-908-664-14	Sequence 14, Appl	939	5	2.3	211	9	US-10-176-925-102	Sequence 102, App
867	5	2.3	203	9	US-10-226-408-3	Sequence 3, Appl	940	5	2.3	211	9	US-10-176-978-102	Sequence 102, App
868	5	2.3	203	9	US-10-177-916-3	Sequence 3, Appl	941	5	2.3	211	9	US-10-179-510-102	Sequence 102, App
869	5	2.3	203	10	US-09-846-729A-5	Sequence 5, Appl	942	5	2.3	211	9	US-10-180-543-102	Sequence 102, App
870	5	2.3	203	10	US-09-846-729A-15	Sequence 15, Appl	943	5	2.3	211	9	US-10-180-544-102	Sequence 102, App
871	5	2.3	203	10	US-09-764-864-1120	Sequence 1120, Ap	944	5	2.3	211	9	US-10-180-546-102	Sequence 102, App
872	5	2.3	204	9	US-09-252-150-15	Sequence 15, Appl	945	5	2.3	211	9	US-10-180-547-102	Sequence 102, App
873	5	2.3	204	12	US-10-116-051-10	Sequence 10, Appl	946	5	2.3	211	9	US-10-180-549-102	Sequence 102, App
874	5	2.3	205	9	US-10-025-367-2	Sequence 2, Appl	947	5	2.3	211	9	US-10-180-555-102	Sequence 102, App
875	5	2.3	205	9	US-09-738-626-3602	Sequence 3602, Ap	948	5	2.3	211	9	US-10-180-559-102	Sequence 102, App
876	5	2.3	205	10	US-09-014-348-11	Sequence 11, Appl	949	5	2.3	211	9	US-10-181-000-102	Sequence 102, App
877	5	2.3	205	10	US-09-156-5	Sequence 5, Appl	950	5	2.3	211	9	US-10-183-010-102	Sequence 102, App
878	5	2.3	205	10	US-09-782-980-41	Sequence 41, Appl	951	5	2.3	211	9	US-10-183-012-102	Sequence 102, App
879	5	2.3	206	9	US-09-873-546-10	Sequence 10, Appl	952	5	2.3	211	9	US-10-184-614-102	Sequence 102, App
880	5	2.3	206	10	US-09-815-242-11988	Sequence 11988, A	953	5	2.3	211	9	US-10-184-623-102	Sequence 102, App
881	5	2.3	208	9	US-09-738-626-6147	Sequence 6147, Ap	954	5	2.3	211	9	US-10-184-635-102	Sequence 102, App
882	5	2.3	208	9	US-10-156-275-72	Sequence 72, Appl	955	5	2.3	211	9	US-10-184-637-102	Sequence 102, App
883	5	2.3	208	10	US-09-925-302-629	Sequence 629, App	956	5	2.3	211	9	US-10-184-646-102	Sequence 102, App
884	5	2.3	208	10	US-09-815-242-11425	Sequence 11425, A	957	5	2.3	211	9	US-10-184-647-102	Sequence 102, App
885	5	2.3	208	10	US-09-027-287-39	Sequence 39, Appl	958	5	2.3	211	9	US-10-184-652-102	Sequence 102, App
886	5	2.3	208	10	US-09-252-6568-39	Sequence 36, Appl	959	5	2.3	211	9	US-10-184-652-102	Sequence 102, App
887	5	2.3	210	10	US-09-765-272-36	Sequence 36, Appl	960	5	2.3	211	9	US-10-187-596-102	Sequence 102, App
888	5	2.3	210	10	US-09-935-390A-36	Sequence 36, Appl	961	5	2.3	211	9	US-10-187-594-102	Sequence 102, App
889	5	2.3	210	10	US-09-731-872-387	Sequence 387, App	962	5	2.3	211	9	US-10-187-745-102	Sequence 102, App
890	5	2.3	211	9	US-09-978-295A-270	Sequence 270, App	963	5	2.3	211	9	US-10-187-885-102	Sequence 102, App
891	5	2.3	211	9	US-09-978-697-270	Sequence 270, App	964	5	2.3	211	9	US-10-187-886-102	Sequence 102, App
892	5	2.3	211	9	US-09-978-192A-270	Sequence 270, App	965	5	2.3	211	9	US-10-189-464-102	Sequence 102, App
893	5	2.3	211	9	US-09-999-832A-270	Sequence 270, App	966	5	2.3	211	9	US-10-176-751-102	Sequence 102, App
894	5	2.3	211	9	US-09-738-626-6578	Sequence 6578, Ap	967	5	2.3	211	9	US-10-176-760-102	Sequence 102, App
895	5	2.3	211	9	US-09-978-189-270	Sequence 270, App	968	5	2.3	211	9	US-10-180-541-102	Sequence 102, App

969 5 2.3 211 9 US-10-180-542-102 Sequence 102, App  
970 5 2.3 211 9 US-10-180-548-102 Sequence 102, App  
971 5 2.3 211 9 US-10-180-551-102 Sequence 102, App  
972 5 2.3 211 9 US-10-180-998-102 Sequence 102, App  
973 5 2.3 211 9 US-10-180-999-102 Sequence 102, App  
974 5 2.3 211 9 US-10-183-013-102 Sequence 102, App  
975 5 2.3 211 9 US-10-184-612-102 Sequence 102, App  
976 5 2.3 211 9 US-10-184-616-102 Sequence 102, App  
977 5 2.3 211 9 US-10-184-617-102 Sequence 102, App  
978 5 2.3 211 9 US-10-184-622-102 Sequence 102, App  
979 5 2.3 211 9 US-10-184-628-102 Sequence 102, App  
980 5 2.3 211 9 US-10-184-629-102 Sequence 102, App  
981 5 2.3 211 9 US-10-184-630-102 Sequence 102, App  
982 5 2.3 211 9 US-10-184-631-102 Sequence 102, App  
983 5 2.3 211 9 US-10-184-632-102 Sequence 102, App  
984 5 2.3 211 9 US-10-184-636-102 Sequence 102, App  
985 5 2.3 211 9 US-10-184-640-102 Sequence 102, App  
986 5 2.3 211 9 US-10-184-650-102 Sequence 102, App  
987 5 2.3 211 9 US-10-184-651-102 Sequence 102, App  
988 5 2.3 211 9 US-10-187-588-102 Sequence 102, App  
989 5 2.3 211 9 US-10-187-597-102 Sequence 102, App  
990 5 2.3 211 9 US-10-187-598-102 Sequence 102, App  
991 5 2.3 211 9 US-10-187-600-102 Sequence 102, App  
992 5 2.3 211 9 US-10-187-601-102 Sequence 102, App  
993 5 2.3 211 9 US-10-187-602-102 Sequence 102, App  
994 5 2.3 211 9 US-10-187-603-102 Sequence 102, App  
995 5 2.3 211 9 US-10-187-741-102 Sequence 102, App  
996 5 2.3 211 9 US-10-187-743-102 Sequence 102, App  
997 5 2.3 211 9 US-10-187-746-102 Sequence 102, App  
998 5 2.3 211 9 US-10-187-747-102 Sequence 102, App  
999 5 2.3 211 9 US-10-187-751-102 Sequence 102, App  
1000 5 2.3 211 9 US-10-187-753-102 Sequence 102, App

## ALIGNMENTS

RESULT 1  
US-10-101-464A-551  
; Sequence 551, Application US/10101464A  
; Publication No. US20030046728A1  
; GENERAL INFORMATION:  
; APPLICANT: Strabala, Timothy  
; APPLICANT: Nieuwenhuizen, Nicolaas  
; APPLICANT: Higgins, Colleen M.  
; TITLE OF INVENTION: Compositions Isolated from Plant Cells  
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling  
; FILE REFERENCE: 11000.1020c2  
; CURRENT APPLICATION NUMBER: US/10/101,464A  
; CURRENT FILING DATE: 2002-03-18  
; PRIOR APPLICATION NUMBER: 09/704,302  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR APPLICATION NUMBER: 09/228,986  
; PRIOR FILING DATE: 1999-01-12  
; PRIOR APPLICATION NUMBER: 60/162,866  
; PRIOR FILING DATE: 1999-11-01  
; PRIOR APPLICATION NUMBER: PCT/US00/00724  
; PRIOR FILING DATE: 2000-01-11  
; NUMBER OF SEQ ID NOS: 989  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 551  
; LENGTH: 151  
; TYPE: PRT  
; ORGANISM: Eucalyptus grandis  
US-10-101-464A-551

Query Match 3.2%; Score 7; DB 9; Length 151;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 39 AKEAIIK 45  
DB 52 AKEAIIK 58

RESULT 2  
US-09-854-133-37  
; Sequence 37, Application US/09854133  
; Publication No. US20020183499A1  
; GENERAL INFORMATION:  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Mohamath, Raedoh  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Secrist, Heather  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.475C10  
; CURRENT APPLICATION NUMBER: US/09/854,133  
; CURRENT FILING DATE: 2001-05-11  
; NUMBER OF SEQ ID NOS: 735  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 37  
; LENGTH: 238  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-854-133-37

Query Match 3.2%; Score 7; DB 9; Length 238;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 170 LNSGSVL 176  
DB 16 LNSGSVL 22

RESULT 3  
US-09-738-973-37  
; Sequence 37, Application US/09738973  
; Patent No. US20020110563A1  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Flinn, Steven P.  
; APPLICANT: Mohamath, Raedoh  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Indrias, Carol Yoseph  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Elliot, Mark  
; APPLICANT: Mannion, Jane  
; APPLICANT: Kalos, Michael D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.475C9  
; CURRENT APPLICATION NUMBER: US/09/738,973  
; CURRENT FILING DATE: 2000-12-14  
; NUMBER OF SEQ ID NOS: 587  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 37  
; LENGTH: 238  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-738-973-37

Query Match 3.2%; Score 7; DB 10; Length 238;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 170 LNSGSVL 176  
DB 16 LNSGSVL 22

RESULT 4  
US-09-854-133-39  
; Sequence 39, Application US/09854133  
; Publication No. US20020183499A1  
; GENERAL INFORMATION:  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Monamath, Raodoh  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Secrist, Heather  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
; FILE REFERENCE: 210121.475C10  
; CURRENT APPLICATION NUMBER: US/09/854,133  
; CURRENT FILING DATE: 2001-05-11  
; NUMBER OF SEQ ID NOS: 735  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 39  
; LENGTH: 243  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-854-133-39

Query Match 3.2%; Score 7; DB 9; Length 243;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 170 LNSGSVL 176  
|||  
Db 30 LNSGSVL 36

RESULT 5  
US-09-854-133-42  
; Sequence 42, Application US/09854133  
; Publication No. US20020183499A1  
; GENERAL INFORMATION:  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Monamath, Raodoh  
; APPLICANT: Benson, Robert A.  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Secrist, Heather  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
; FILE REFERENCE: 210121.475C10  
; CURRENT APPLICATION NUMBER: US/09/854,133  
; CURRENT FILING DATE: 2001-05-11  
; NUMBER OF SEQ ID NOS: 735  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 42  
; LENGTH: 243  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-854-133-42

Query Match 3.2%; Score 7; DB 9; Length 243;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 170 LNSGSVL 176  
|||  
Db 21 LNSGSVL 27

RESULT 6  
US-09-738-973-39  
; Sequence 39, Application US/09738973  
; Patent No. US20020110563A1  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Fling, Steven P.

; APPLICANT: Monamath, Raodoh  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Indrias, Carol Yoseph  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Elliot, Mark  
; APPLICANT: Mannion, Jane  
; APPLICANT: Kalos, Michael D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
; FILE REFERENCE: 210121.475C9  
; CURRENT APPLICATION NUMBER: US/09/738,973  
; CURRENT FILING DATE: 2000-12-14  
; NUMBER OF SEQ ID NOS: 587  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 39  
; LENGTH: 243  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-738-973-39

Query Match 3.2%; Score 7; DB 10; Length 243;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 170 LNSGSVL 176  
|||  
Db 30 LNSGSVL 36

RESULT 7  
US-09-738-973-42  
; Sequence 42, Application US/09738973  
; Patent No. US20020110563A1  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Fling, Steven P.  
; APPLICANT: Monamath, Raodoh  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Indrias, Carol Yoseph  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Elliot, Mark  
; APPLICANT: Mannion, Jane  
; APPLICANT: Kalos, Michael D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
; FILE REFERENCE: 210121.475C9  
; CURRENT APPLICATION NUMBER: US/09/738,973  
; CURRENT FILING DATE: 2000-12-14  
; NUMBER OF SEQ ID NOS: 587  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 42  
; LENGTH: 243  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-738-973-42

Query Match 3.2%; Score 7; DB 10; Length 243;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 170 LNSGSVL 176  
|||  
Db 21 LNSGSVL 27

RESULT 8  
US-09-854-133-43  
; Sequence 43, Application US/09854133  
; Publication No. US20020183499A1

```
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 43
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-854-133-43

Query Match          3.2% Score 7; DB 9; Length 244;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 LNSGSVL 176
Db 22 LNSGSVL 28

RESULT 9
US-09-854-133-46
; Sequence 46, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 46
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-854-133-46

Query Match          3.2% Score 7; DB 9; Length 244;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 LNSGSVL 176
Db 22 LNSGSVL 28

RESULT 10
US-09-738-973-43
; Sequence 43, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Fling, Steven P.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Algate, Paul A.
; APPLICANT: Secrist, Heather
; APPLICANT: Indrias, Carol Yoseph
```

```
; APPLICANT: Benson, Darin R.
; APPLICANT: Elliott, Mark
; APPLICANT: Mannion, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 210121.475C9
; CURRENT APPLICATION NUMBER: US/09/738,973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 43
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-738-973-43

Query Match          3.2% Score 7; DB 10; Length 244;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 LNSGSVL 176
Db 22 LNSGSVL 28

RESULT 11
US-09-738-973-46
; Sequence 46, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Fling, Steven P.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Algate, Paul A.
; APPLICANT: Secrist, Heather
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Elliott, Mark
; APPLICANT: Mannion, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 210121.475C9
; CURRENT APPLICATION NUMBER: US/09/738,973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 46
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-738-973-46

Query Match          3.2% Score 7; DB 10; Length 244;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 LNSGSVL 176
Db 22 LNSGSVL 28

RESULT 12
US-09-854-133-40
; Sequence 40, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Henderson, Robert A.
```

```

; APPLICANT: Benson, Darin R.
; APPLICANT: Secret, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; FILE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854.133
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 40
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-854-133-40
```

```
Query Match          3.2%; Score 7; DB 9; Length 245;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 170 LNSGSVL 176
      |||||
Db 23 LNSGSVL 29
```

```

RESULT 13
US-09-738-973-40
; Sequence 40, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Flinn, Steven P.
; APPLICANT: Mohamath, Raedoh
; APPLICANT: Algate, Paul A.
; APPLICANT: Secret, Heather
; APPLICANT: Indirias, Carol Joseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Elliot, Mark
; APPLICANT: Mannion, Jane
; APPLICANT: Kajos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; FILE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C9
; CURRENT APPLICATION NUMBER: US/09/738.973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 40
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-738-973-40
```

```
Query Match          3.2%; Score 7; DB 10; Length 245;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 170 LNSGSVL 176
      |||||
Db 23 LNSGSVL 29
```

```

RESULT 14
US-09-815-242-11873
; Sequence 11873, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
```

```

; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: EITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815.242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11873
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-11873
```

```
Query Match          3.2%; Score 7; DB 10; Length 311;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 116 QAAVILL 122
      |||||
Db 63 QAAVILL 69
```

```

RESULT 15
US-09-773-882-4
; Sequence 4, Application US/09773882
; Patent No. US20020106769A1
; GENERAL INFORMATION:
; APPLICANT: Omura, Mitsuo
; APPLICANT: Matsumoto, Ryoji
; APPLICANT: Moriguchi, Takaya
; APPLICANT: Haegawa, Shin
; APPLICANT: Subayda, Charles
; TITLE OF INVENTION: UDP-D-Glucose:limonoid Glucosyltransferase
; FILE REFERENCE: 0119.98
; CURRENT APPLICATION NUMBER: US/09/773.882
; CURRENT FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: JP 2000-021179
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 4
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Citrus unshiu
US-09-773-882-4
```

```
Query Match          2.8%; Score 6; DB 10; Length 14;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 97 TPVGDG 102
      |||||
Db 9 TPVGDG 14
```



```
RESULT 16
US-09-976-674-61
; Sequence 61, Application US/09976674
; Patent No. US2002015843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinaanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junier, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPP1
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-61

Query Match      2.8%; Score 6; DB 10; Length 15;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 KEATLK 45
   |||||
Db 1 KEATLK 6

RESULT 17
US-09-864-761-41312
; Sequence 41312, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
```

```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 41312
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL136306.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN EST_HUMAN HIT: AA315245.1, EVALU6 7.00e-09
US-09-864-761-41312

Query Match      2.8%; Score 6; DB 10; Length 28;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 210 VLRGVN 215
   |||||
Db 8 VLRGVN 13

RESULT 18
US-09-864-761-36602
; Sequence 36602, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
```

```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 36602
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL022725.8
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.67
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.96
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.85
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.38
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.73
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.79
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.96
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.99
; OTHER INFORMATION: SWISSPROT HIT: P46448, EVALUATE 2.10e+00
US-09-864-761-36602
```

```
Query Match      2.8%; Score 6; DB 10; Length 34;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      142 QTNPNV 147
        |||||
Db       24 QTNPNV 29
```

```
RESULT 19
US-09-864-761-33862
; Sequence 33862, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
```

```
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33862
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007670.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.79
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.84
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.94
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.75
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.8
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 8.1
; OTHER INFORMATION: EST HUMAN HIT: AU130635.1, EVALUATE 1.00e-17
; OTHER INFORMATION: SWISSPROT HIT: P87136, EVALUATE 8.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: BS397236.1, EVALUATE 9.00e-18
US-09-864-761-33862
```

```
Query Match      2.8%; Score 6; DB 10; Length 43;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      205 TEEVVV 210
        |||||
Db       23 TEEVVV 28
```

```
RESULT 20
US-09-981-876-242
; Sequence 242, Application US/09981876
; Patent No. US2002016469A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: P2001P1
; CURRENT APPLICATION NUMBER: US/09/981,876
; CURRENT FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: 09/148,545
; PRIOR FILING DATE: 1998-09-04
; PRIOR APPLICATION NUMBER: 60/040,162
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,333
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/038,621
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,161
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,626
```

1	PRIOR FILING DATE: 1997-03-07	
2	PRIOR APPLICATION NUMBER: 60/040,334	
3	PRIOR FILING DATE: 1997-03-07	
4	PRIOR APPLICATION NUMBER: 60/040,336	
5	PRIOR FILING DATE: 1997-03-07	
6	PRIOR APPLICATION NUMBER: 60/040,163	
7	PRIOR FILING DATE: 1997-03-07	
8	PRIOR APPLICATION NUMBER: 60/047,615	
9	PRIOR FILING DATE: 1997-05-23	
10	PRIOR APPLICATION NUMBER: 60/047,600	
11	PRIOR FILING DATE: 1997-05-23	
12	PRIOR APPLICATION NUMBER: 60/047,597	
13	PRIOR FILING DATE: 1997-05-23	
14	PRIOR APPLICATION NUMBER: 60/047,502	
15	PRIOR FILING DATE: 1997-05-23	
16	PRIOR APPLICATION NUMBER: 60/047,633	
17	PRIOR FILING DATE: 1997-05-23	
18	PRIOR APPLICATION NUMBER: 60/047,583	
19	PRIOR FILING DATE: 1997-05-23	
20	PRIOR APPLICATION NUMBER: 60/047,617	
21	PRIOR FILING DATE: 1997-05-23	
22	PRIOR APPLICATION NUMBER: 60/047,618	
23	PRIOR FILING DATE: 1997-05-23	
24	PRIOR APPLICATION NUMBER: 60/047,503	
25	PRIOR FILING DATE: 1997-05-23	
26	PRIOR APPLICATION NUMBER: 60/047,592	
27	PRIOR FILING DATE: 1997-05-23	
28	PRIOR APPLICATION NUMBER: 60/047,581	
29	PRIOR FILING DATE: 1997-05-23	
30	PRIOR APPLICATION NUMBER: 60/047,584	
31	PRIOR FILING DATE: 1997-05-23	
32	PRIOR APPLICATION NUMBER: 60/047,500	
33	PRIOR FILING DATE: 1997-05-23	
34	PRIOR APPLICATION NUMBER: 60/047,587	
35	PRIOR FILING DATE: 1997-05-23	
36	PRIOR APPLICATION NUMBER: 60/047,492	
37	PRIOR FILING DATE: 1997-05-23	
38	PRIOR APPLICATION NUMBER: 60/047,598	
39	PRIOR FILING DATE: 1997-05-23	
40	PRIOR APPLICATION NUMBER: 60/047,613	
41	PRIOR FILING DATE: 1997-05-23	
42	PRIOR APPLICATION NUMBER: 60/047,582	
43	PRIOR FILING DATE: 1997-05-23	
44	PRIOR APPLICATION NUMBER: 60/047,596	
45	PRIOR FILING DATE: 1997-05-23	
46	PRIOR APPLICATION NUMBER: 60/047,612	
47	PRIOR FILING DATE: 1997-05-23	
48	PRIOR APPLICATION NUMBER: 60/047,632	
49	PRIOR FILING DATE: 1997-05-23	
50	PRIOR APPLICATION NUMBER: 60/047,601	
51	PRIOR FILING DATE: 1997-05-23	
52	PRIOR APPLICATION NUMBER: 60/043,580	
53	PRIOR FILING DATE: 1997-04-11	
54	PRIOR APPLICATION NUMBER: 60/043,568	
55	PRIOR FILING DATE: 1997-04-11	
56	PRIOR APPLICATION NUMBER: 60/043,314	
57	PRIOR FILING DATE: 1997-04-11	
58	PRIOR APPLICATION NUMBER: 60/043,569	
59	PRIOR FILING DATE: 1997-04-11	
60	PRIOR APPLICATION NUMBER: 60/043,311	
61	PRIOR FILING DATE: 1997-04-11	
62	PRIOR APPLICATION NUMBER: 60/043,671	
63	PRIOR FILING DATE: 1997-04-11	
64	PRIOR APPLICATION NUMBER: 60/043,674	
65	PRIOR FILING DATE: 1997-04-11	
66	PRIOR APPLICATION NUMBER: 60/043,669	
67	PRIOR FILING DATE: 1997-04-11	
68	PRIOR APPLICATION NUMBER: 60/043,312	
69	PRIOR FILING DATE: 1997-04-11	
70	PRIOR APPLICATION NUMBER: 60/043,313	
71	PRIOR FILING DATE: 1997-04-11	
72	PRIOR APPLICATION NUMBER: 60/043,672	
73	PRIOR FILING DATE: 1997-04-11	

PRIOR APPLICATION NUMBER: 60/043,315	PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: 60/048,974	PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/056,886	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,877	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,889	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,893	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,630	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,878	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,662	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,872	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,882	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,637	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,903	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,888	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,879	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,880	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,894	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,911	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,636	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,874	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,910	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,864	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,631	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,845	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,892	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/047,595	PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/057,761	PRIOR FILING DATE: 05-Sep-1997
PRIOR APPLICATION NUMBER: 60/047,599	PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,588	PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,585	PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,586	PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,590	PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,594	PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,589	PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,593	PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,614	PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/043,578	PRIOR FILING DATE: 1997-05-23

;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,576  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/047,501  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/043,670  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/056,632  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,664  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,876  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,881  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,909  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,875  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,862  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,887  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,908  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/048,964  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/057,650  
;; PRIOR FILING DATE: 1997-09-05  
;; PRIOR APPLICATION NUMBER: 60/056,884  
;; PRIOR FILING DATE: 1997-08-22  
;; NUMBER OF SEQ ID NOS: 280  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 242  
;; LENGTH: 58

Query Match 2.8%; Score 6; DB 9; Length 58;  
Best Local Similarity 100.0%; Pred. No. 96;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 172 SGSUV 177  
Db 50 SGSUV 55

## RESULT 21

US-09-148-545-242  
;; Sequence 242, Application US/09148545  
;; Publication No. US20030027132a1  
;; GENERAL INFORMATION:  
;; APPLICANT: Rosen et al.  
;; TITLE OF INVENTION: 70 Human Secreted Proteins  
;; FILE REFERENCE: P2001P1  
;; CURRENT APPLICATION NUMBER: US/09/148,545  
;; CURRENT FILING DATE: 1998-09-04  
;; EARLIER APPLICATION NUMBER: PCT/US98/04482  
;; EARLIER FILING DATE: 1998-03-06  
;; EARLIER APPLICATION NUMBER: 60/040,162  
;; EARLIER FILING DATE: 1997-03-07  
;; EARLIER APPLICATION NUMBER: 60/040,333  
;; EARLIER FILING DATE: 1997-03-07  
;; EARLIER APPLICATION NUMBER: 60/038,621  
;; EARLIER FILING DATE: 1997-03-07  
;; EARLIER APPLICATION NUMBER: 60/040,161  
;; EARLIER FILING DATE: 1997-03-07  
;; EARLIER APPLICATION NUMBER: 60/040,626  
;; EARLIER FILING DATE: 1997-03-07  
;; EARLIER APPLICATION NUMBER: 60/040,334  
;; EARLIER FILING DATE: 1997-03-07  
;; EARLIER APPLICATION NUMBER: 60/040,336  
;; EARLIER FILING DATE: 1997-03-07  
;; EARLIER APPLICATION NUMBER: 60/040,163  
;; EARLIER FILING DATE: 1997-03-07

;; EARLIER APPLICATION NUMBER: 60/047,615  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,600  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,597  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,502  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,633  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,583  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,617  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,618  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,503  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,592  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,581  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,584  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,500  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,587  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,492  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,598  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,613  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,582  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,596  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,612  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,632  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,601  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/043,580  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/043,568  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/043,314  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/043,569  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/043,311  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/043,671  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/043,674  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/043,669  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/043,312  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/043,313  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/043,672  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/043,315  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/048,974  
;; EARLIER FILING DATE: 1997-06-06  
;; EARLIER APPLICATION NUMBER: 60/056,886  
;; EARLIER FILING DATE: 1997-08-22  
;; EARLIER APPLICATION NUMBER: 60/056,877

EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 889  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 893  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 630  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 878  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 662  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 872  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 882  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 637  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 903  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 888  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 879  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 880  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 894  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 911  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 636  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 874  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 910  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 864  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 631  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 845  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 892  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/047, 595  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/057, 761  
EARLIER FILING DATE: 05-Sep-1997  
EARLIER APPLICATION NUMBER: 60/047, 599  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 588  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 585  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 586  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 590  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 594  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 589  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 593  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 614  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043, 578  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043, 576  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/047, 501  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043, 670  
EARLIER FILING DATE: 1997-04-11

EARLIER APPLICATION NUMBER: 60/056, 632  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 664  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 876  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 881  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 909  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 875  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 862  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 887  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 908  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/048, 964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/057, 650  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/056, 884  
EARLIER FILING DATE: 1997-08-22  
NUMBER OF SEQ ID NOS: 280  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 242  
LENGTH: 58

Query Match 2.8%; Score 6; DB 9; Length 58;  
Best Local Similarity 100.0%; Fred. No. 96;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 172 SGGVLY 177  
Db 50 SGGVLY 55

RESULT 22  
US-09-981-876-186  
Sequence 186; Application US/09981876  
Patent No. US2002016669A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 70 Human Secreted Proteins  
FILE REFERENCE: P2001P1  
CURRENT APPLICATION NUMBER: US/09/981,876  
CURRENT FILING DATE: 2001-10-19  
PRIOR APPLICATION NUMBER: 09/148,545  
PRIOR FILING DATE: 1998-09-04  
PRIOR APPLICATION NUMBER: 60/040,162  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/040,333  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/038,621  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/040,161  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/040,626  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/040,334  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/040,336  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/040,163  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/047,615  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,600  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,597  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,502

PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,633
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,583
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,617
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,618
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,503
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,592
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,581
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,584
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,500
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,587
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,492
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,598
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,613
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,582
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,596
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,612
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/043,580
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: 60/043,568
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: 60/043,314
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: 60/043,569
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: 60/043,311
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: 60/043,671
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: 60/043,674
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: 60/043,669
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: 60/043,312
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: 60/043,313
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: 60/043,672
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: 60/043,315
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: 60/048,974
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/056,886
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,877
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,889
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,893
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,630
PRIOR FILING DATE: 1997-08-22

1	PRIOR APPLICATION NUMBER: 60/056, 878
2	PRIOR FILING DATE: 1997-08-22
3	PRIOR APPLICATION NUMBER: 60/056, 662
4	PRIOR FILING DATE: 1997-08-22
5	PRIOR APPLICATION NUMBER: 60/056, 872
6	PRIOR FILING DATE: 1997-08-22
7	PRIOR APPLICATION NUMBER: 60/056, 882
8	PRIOR FILING DATE: 1997-08-22
9	PRIOR APPLICATION NUMBER: 60/056, 637
10	PRIOR FILING DATE: 1997-08-22
11	PRIOR APPLICATION NUMBER: 60/056, 903
12	PRIOR FILING DATE: 1997-08-22
13	PRIOR APPLICATION NUMBER: 60/056, 888
14	PRIOR FILING DATE: 1997-08-22
15	PRIOR APPLICATION NUMBER: 60/056, 879
16	PRIOR FILING DATE: 1997-08-22
17	PRIOR APPLICATION NUMBER: 60/056, 880
18	PRIOR FILING DATE: 1997-08-22
19	PRIOR APPLICATION NUMBER: 60/056, 894
20	PRIOR FILING DATE: 1997-08-22
21	PRIOR APPLICATION NUMBER: 60/056, 911
22	PRIOR FILING DATE: 1997-08-22
23	PRIOR APPLICATION NUMBER: 60/056, 636
24	PRIOR FILING DATE: 1997-08-22
25	PRIOR APPLICATION NUMBER: 60/056, 874
26	PRIOR FILING DATE: 1997-08-22
27	PRIOR APPLICATION NUMBER: 60/056, 910
28	PRIOR FILING DATE: 1997-08-22
29	PRIOR APPLICATION NUMBER: 60/056, 864
30	PRIOR FILING DATE: 1997-08-22
31	PRIOR APPLICATION NUMBER: 60/056, 631
32	PRIOR FILING DATE: 1997-08-22
33	PRIOR APPLICATION NUMBER: 60/056, 845
34	PRIOR FILING DATE: 1997-08-22
35	PRIOR APPLICATION NUMBER: 60/056, 892
36	PRIOR FILING DATE: 1997-08-22
37	PRIOR APPLICATION NUMBER: 60/047, 595
38	PRIOR FILING DATE: 1997-05-23
39	PRIOR APPLICATION NUMBER: 60/057, 761
40	PRIOR FILING DATE: 05-Sep-1997
41	PRIOR APPLICATION NUMBER: 60/047, 599
42	PRIOR FILING DATE: 1997-05-23
43	PRIOR APPLICATION NUMBER: 60/047, 588
44	PRIOR FILING DATE: 1997-05-23
45	PRIOR APPLICATION NUMBER: 60/047, 585
46	PRIOR FILING DATE: 1997-05-23
47	PRIOR APPLICATION NUMBER: 60/047, 586
48	PRIOR FILING DATE: 1997-05-23
49	PRIOR APPLICATION NUMBER: 60/047, 590
50	PRIOR FILING DATE: 1997-05-23
51	PRIOR APPLICATION NUMBER: 60/047, 594
52	PRIOR FILING DATE: 1997-05-23
53	PRIOR APPLICATION NUMBER: 60/047, 589
54	PRIOR FILING DATE: 1997-04-11
55	PRIOR APPLICATION NUMBER: 60/047, 593
56	PRIOR FILING DATE: 1997-04-11
57	PRIOR APPLICATION NUMBER: 60/047, 614
58	PRIOR FILING DATE: 1997-05-23
59	PRIOR APPLICATION NUMBER: 60/043, 578
60	PRIOR FILING DATE: 1997-04-11
61	PRIOR APPLICATION NUMBER: 60/043, 576
62	PRIOR FILING DATE: 1997-04-11
63	PRIOR APPLICATION NUMBER: 60/047, 501
64	PRIOR FILING DATE: 1997-05-23
65	PRIOR APPLICATION NUMBER: 60/043, 670
66	PRIOR FILING DATE: 1997-04-11
67	PRIOR APPLICATION NUMBER: 60/056, 632
68	PRIOR FILING DATE: 1997-08-22
69	PRIOR APPLICATION NUMBER: 60/056, 664
70	PRIOR FILING DATE: 1997-08-22
71	PRIOR APPLICATION NUMBER: 60/056, 876
72	PRIOR FILING DATE: 1997-08-22
73	PRIOR APPLICATION NUMBER: 60/056, 881

```

; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,909
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,875
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,862
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,887
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,908
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/048,964
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/057,650
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/056,884
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 280
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 186
; LENGTH: 59

Query Match      2.8%; Score 6; DB 9; Length 59;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 172 SCSVLV 177
Db 50 SCSVLV 55

RESULT 23
US-09-148-545-186
; Sequence 186, Application US/09148545
; Publication No. US20030027132A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: P2001P1
; CURRENT APPLICATION NUMBER: US/09/148,545
; CURRENT FILING DATE: 1998-09-04
; EARLIER APPLICATION NUMBER: PCT/US98/04482
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,161
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
```

1	EARLIER	FILING DATE:	1997-08-22	60/056, 637
2	EARLIER	FILING DATE:	1997-08-22	60/056, 637
3	EARLIER	FILING DATE:	1997-08-22	60/056, 903
4	EARLIER	FILING DATE:	1997-08-22	60/056, 903
5	EARLIER	FILING DATE:	1997-08-22	60/056, 888
6	EARLIER	FILING DATE:	1997-08-22	60/056, 888
7	EARLIER	FILING DATE:	1997-08-22	60/056, 894
8	EARLIER	FILING DATE:	1997-08-22	60/056, 894
9	EARLIER	FILING DATE:	1997-08-22	60/056, 911
10	EARLIER	FILING DATE:	1997-08-22	60/056, 911
11	EARLIER	FILING DATE:	1997-08-22	60/056, 910
12	EARLIER	FILING DATE:	1997-08-22	60/056, 910
13	EARLIER	FILING DATE:	1997-08-22	60/056, 864
14	EARLIER	FILING DATE:	1997-08-22	60/056, 864
15	EARLIER	FILING DATE:	1997-08-22	60/056, 631
16	EARLIER	FILING DATE:	1997-08-22	60/056, 631
17	EARLIER	FILING DATE:	1997-08-22	60/056, 845
18	EARLIER	FILING DATE:	1997-08-22	60/056, 845
19	EARLIER	FILING DATE:	1997-08-22	60/056, 892
20	EARLIER	FILING DATE:	1997-08-22	60/056, 892
21	EARLIER	FILING DATE:	1997-08-22	60/047, 595
22	EARLIER	FILING DATE:	1997-05-23	60/047, 595
23	EARLIER	FILING DATE:	1997-05-23	60/057, 761
24	EARLIER	FILING DATE:	05-Sep-1997	60/057, 761
25	EARLIER	FILING DATE:	05-Sep-1997	60/047, 599
26	EARLIER	FILING DATE:	1997-05-23	60/047, 599
27	EARLIER	FILING DATE:	1997-05-23	60/047, 588
28	EARLIER	FILING DATE:	1997-05-23	60/047, 588
29	EARLIER	FILING DATE:	1997-05-23	60/047, 585
30	EARLIER	FILING DATE:	1997-05-23	60/047, 585
31	EARLIER	FILING DATE:	1997-05-23	60/047, 586
32	EARLIER	FILING DATE:	1997-05-23	60/047, 586
33	EARLIER	FILING DATE:	1997-05-23	60/047, 590
34	EARLIER	FILING DATE:	1997-05-23	60/047, 590
35	EARLIER	FILING DATE:	1997-05-23	60/047, 594
36	EARLIER	FILING DATE:	1997-05-23	60/047, 594
37	EARLIER	FILING DATE:	1997-05-23	60/047, 589
38	EARLIER	FILING DATE:	1997-05-23	60/047, 589
39	EARLIER	FILING DATE:	1997-05-23	60/043, 578
40	EARLIER	FILING DATE:	1997-04-11	60/043, 578
41	EARLIER	FILING DATE:	1997-04-11	60/043, 576
42	EARLIER	FILING DATE:	1997-04-11	60/043, 576
43	EARLIER	FILING DATE:	1997-04-11	60/047, 501
44	EARLIER	FILING DATE:	1997-05-23	60/043, 670
45	EARLIER	FILING DATE:	1997-04-11	60/043, 670
46	EARLIER	FILING DATE:	1997-04-11	60/056, 632
47	EARLIER	FILING DATE:	1997-08-22	60/056, 632
48	EARLIER	FILING DATE:	1997-08-22	60/056, 664
49	EARLIER	FILING DATE:	1997-08-22	60/056, 664
50	EARLIER	FILING DATE:	1997-08-22	60/056, 876
51	EARLIER	FILING DATE:	1997-08-22	60/056, 881
52	EARLIER	FILING DATE:	1997-08-22	60/056, 881
53	EARLIER	FILING DATE:	1997-08-22	60/056, 909
54	EARLIER	FILING DATE:	1997-08-22	60/056, 909
55	EARLIER	FILING DATE:	1997-08-22	60/056, 875
56	EARLIER	FILING DATE:	1997-08-22	60/056, 875
57	EARLIER	FILING DATE:	1997-08-22	60/056, 862
58	EARLIER	FILING DATE:	1997-08-22	60/056, 862

```

; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 280
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 186
; LENGTH: 59
;
Query Match          2.8%; Score 6; DB 9; Length 59;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      172 GGSVLV 177
      |||||
Db      50 GGSVLV 55

RESULT 24
US-09-864-761-45366
; Sequence 45366, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecmics-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203

```



```

; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 45366
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL121914.16
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.86
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.49
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.72
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.88
; OTHER INFORMATION: EST HUMAN HIT: R28249.1, EVALUATE 4.00e-28
; OTHER INFORMATION: SWISSPROT HIT: O21326, EVALUATE 9.00e-01
; US-09-864-761-45366

Query Match      2.8%; Score 6; DB 10; Length 62;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 42 AILKIN 47
    |||||
Db 40 AILKIN 45

RESULT 25
; US-09-864-761-46027
; Sequence 46027, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
```

```

; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 46027
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL138828.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.1
; OTHER INFORMATION: EXPRESSED IN SWISSPROT HIT: P40497, EVALUATE 3.40e+00
; OTHER INFORMATION: SWISSPROT HIT: A1087936.1, EVALUATE 4.00e-23
; US-09-864-761-46027

Query Match      2.8%; Score 6; DB 10; Length 70;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 206 EEVVL 211
    |||||
Db 15 EEVVL 20

RESULT 26
; US-09-738-626-6744
; Sequence 6744, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: YATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6744
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-09-738-626-6744

Query Match      2.8%; Score 6; DB 9; Length 87;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 118 AVILLT 123
    |||||
Db 73 AVILLT 78

RESULT 27
```

US-09-984-271-178  
; Sequence 178, Application US/09984271  
; Publication No. US20030040088A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 71 Human Secreted Proteins  
; FILE REFERENCE: PZ030P1  
; CURRENT APPLICATION NUMBER: US/09/984,271  
; PRIOR FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: 09/482,273  
; PRIOR FILING DATE: 2000-01-13  
; PRIOR APPLICATION NUMBER: PCT/US99/15849  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: 60/092,921  
; PRIOR FILING DATE: 1998-07-15  
; PRIOR APPLICATION NUMBER: 60/092,922  
; PRIOR FILING DATE: 1998-07-15  
; PRIOR APPLICATION NUMBER: 60/092,956  
; PRIOR FILING DATE: 1998-07-15  
; NUMBER OF SEQ ID NOS: 267  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 178  
; LENGTH: 88  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (88)  
; OTHER INFORMATION: Xaa equals stop translation  
US-09-984-271-178

Query Match 2.8%; Score 6; DB 9; Length 88;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 117 AAATLL 122  
|||  
Db 18 AAATLL 23

RESULT 28  
US-09-984-271-179  
; Sequence 179, Application US/09984271  
; Publication No. US20030040088A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 71 Human Secreted Proteins  
; FILE REFERENCE: PZ030P1  
; CURRENT APPLICATION NUMBER: US/09/984,271  
; PRIOR FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: 09/482,273  
; PRIOR FILING DATE: 2000-01-13  
; PRIOR APPLICATION NUMBER: PCT/US99/15849  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: 60/092,921  
; PRIOR FILING DATE: 1998-07-15  
; PRIOR APPLICATION NUMBER: 60/092,922  
; PRIOR FILING DATE: 1998-07-15  
; PRIOR APPLICATION NUMBER: 60/092,956  
; PRIOR FILING DATE: 1998-07-15  
; NUMBER OF SEQ ID NOS: 267  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 179  
; LENGTH: 88  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (55)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (66)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE  
; LOCATION: (88)  
; OTHER INFORMATION: Xaa equals stop translation  
US-09-984-271-179

Query Match 2.8%; Score 6; DB 9; Length 88;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 117 AAATLL 122  
|||  
Db 18 AAATLL 23

RESULT 29  
US-09-864-761-37228  
; Sequence 37228, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aeomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; PRIOR FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 37228  
; LENGTH: 93  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL021877.1  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7

```
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.6
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.1
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.2
OTHER INFORMATION: SWISSPROT HIT: P05499, EVALUATE 4.50e+00
OTHER INFORMATION: EST_HUMAN HIT: BF347445.1, EVALUATE 9.00e-26
US-09-864-761-37228
```

```
Query Match 2.8%; Score 6; DB 10; Length 93;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3 VTASCT 8
    |||||
Db 12 VTASCT 17
```

```
RESULT 30
US-10-012-896-933
Sequence 933, Application US/10012896
Publication No. US20020183251A1
GENERAL INFORMATION:
APPLICANT: Xu, Jianshun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yugu
APPLICANT: Kalos, Michael D.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darriek
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Hural, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals de Bassols, Carlota
APPLICANT: Foy, Teresa
APPLICANT: Fanger, Gary R.
APPLICANT: Mantanabe, Yoshihiro
APPLICANT: Meagher, Madeline Joy
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C27
CURRENT APPLICATION NUMBER: US/10/012, 896
CURRENT FILING DATE: 2001-12-10
NUMBER OF SEQ ID NOS: 1011
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 933
LENGTH: 100
TYPE: PRT
ORGANISM: Homo sapiens
US-10-012-896-933
```

```
Query Match 2.8%; Score 6; DB 9; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 169 AINSGS 174
    |||||
Db 2 AINSGS 7
```

```
RESULT 31
US-09-895-793-933
Sequence 933, Application US/09895793
Publication No. US20020192763A1
GENERAL INFORMATION:
```

```
APPLICANT: Xu, Jianshun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yugu
APPLICANT: Kalos, Michael D.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darriek
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Hural, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals de Bassols, Carlota
APPLICANT: Foy, Teresa
APPLICANT: Fanger, Gary R.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.534C2
CURRENT APPLICATION NUMBER: US/09/895, 793
CURRENT FILING DATE: 2001-06-29
NUMBER OF SEQ ID NOS: 982
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 933
LENGTH: 100
TYPE: PRT
ORGANISM: Homo sapiens
US-09-895-793-933
```

```
Query Match 2.8%; Score 6; DB 9; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 169 AINSGS 174
    |||||
Db 2 AINSGS 7
```

```
RESULT 32
US-09-895-814-933
Sequence 933, Application US/09895814
Publication No. US20020193296A1
GENERAL INFORMATION:
APPLICANT: Xu, Jianshun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yugu
APPLICANT: Kalos, Michael D.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darriek
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Hural, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
```

```

; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895.814
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 933
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-895-814-933

Query Match      2.8%; Score 6; DB 9; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 ALNSGS 174
Db 2 ALNSGS 7

RESULT 33
US-09-759-143-933
; Sequence 933, Application US/09759143
; Patent No. US20020022248A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kaios, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolck, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Derrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C23
; CURRENT APPLICATION NUMBER: US/09/759.143
; CURRENT FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 933
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-759-143-933

Query Match      2.8%; Score 6; DB 10; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 ALNSGS 174
Db 2 ALNSGS 7

RESULT 34
US-09-780-669-933
; Sequence 933, Application US/09780669
; Patent No. US20020051977A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kaios, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolck, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Derrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C24
; CURRENT APPLICATION NUMBER: US/09/780.669
; CURRENT FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 943
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 933
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-780-669-933

Query Match      2.8%; Score 6; DB 10; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 ALNSGS 174
Db 2 ALNSGS 7

RESULT 35
US-09-822-827-933
; Sequence 933, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822.827
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 933
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-822-827-933

Query Match      2.8%; Score 6; DB 10; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 ALNSGS 174
Db 2 ALNSGS 7

RESULT 36
US-09-796-692-1175
```

```
Sequence 1175, Application US/09796692
Publication No. US2002019862A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
FILE REFERENCE: 2077.001200
CURRENT APPLICATION NUMBER: US/09/796,692
CURRENT FILING DATE: 2001-03-01
PRIOR APPLICATION NUMBER: 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: 60/223,378
PRIOR FILING DATE: 2000-08-07
NUMBER OF SEQ ID NOS: 9597
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1175
LENGTH: 109
TYPE: PRT
ORGANISM: Homo sapiens
US-09-796-692-1175

Query Match      2.8%: Score 6; DB 9; Length 109;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 VVLRGV 214
DB 47 VVLRGV 52

RESULT 37
US-09-957-708-30
Sequence 30, Application US/09957708
Publication No. US20030031678A1
GENERAL INFORMATION:
APPLICANT: Sun, Yongming
APPLICANT: Recipon, Herve
APPLICANT: Caferkey, Robert
APPLICANT: Ali, Shujath
TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific
FILE REFERENCE: DEX-0239
CURRENT APPLICATION NUMBER: US/09/957,708
CURRENT FILING DATE: 2001-09-19
PRIOR APPLICATION NUMBER: 60/233,746
PRIOR FILING DATE: 2000-09-19
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 30
LENGTH: 122
TYPE: PRT
```

```
ORGANISM: Homo sapiens
US-09-957-708-30

Query Match      2.8%: Score 6; DB 9; Length 122;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 PVEEYK 138
DB 38 PVEEYK 43

RESULT 38
US-09-738-626-5132
Sequence 5132, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 5132
LENGTH: 132
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-5132

Query Match      2.8%: Score 6; DB 9; Length 132;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 FADDAE 159
DB 44 FADDAE 49

RESULT 39
US-09-764-853-560
Sequence 560, Application US/09764853
Patent No. US20020090672A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PJ206
CURRENT APPLICATION NUMBER: US/09/764,853
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 939
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 560
LENGTH: 139
TYPE: PRT
ORGANISM: Homo sapiens
US-09-764-853-560
```

Query Match 2.8%; Score 6; DB 10; Length 139;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 116 QAAVIL 121  
 |||||  
 DB 69 QAAVIL 74

RESULT 40

US-09-815-242-13435  
 ; Sequence 13435, Application US/09815242  
 ; Patent No. US20020061569A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Haselbeck, Robert  
 ; APPLICANT: Ohlsen, Karl L.  
 ; APPLICANT: Zyekind, Judith W.  
 ; APPLICANT: Wall, Daniel  
 ; APPLICANT: Trawick, John D.  
 ; APPLICANT: Carr, Grant J.  
 ; APPLICANT: Yamamoto, Robert T.  
 ; APPLICANT: Xu, H. Howard  
 ; TITLE OR INVENTION: Identification of Essential Genes in  
 ; FILE REFERENCE: ELITRA.011A  
 ; CURRENT APPLICATION NUMBER: US/09/815,242  
 ; CURRENT FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: 60/191,078  
 ; PRIOR FILING DATE: 2000-03-21  
 ; PRIOR APPLICATION NUMBER: 60/206,848  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/207,727  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: 60/242,578  
 ; PRIOR FILING DATE: 2000-10-23  
 ; PRIOR APPLICATION NUMBER: 60/253,625  
 ; PRIOR FILING DATE: 2000-11-27  
 ; PRIOR APPLICATION NUMBER: 60/257,931  
 ; PRIOR FILING DATE: 2000-12-22  
 ; PRIOR APPLICATION NUMBER: 60/269,308  
 ; PRIOR FILING DATE: 2001-02-16  
 ; NUMBER OF SEQ ID NOS: 14110  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 13435  
 ; LENGTH: 150  
 ; TYPE: PRT  
 ; ORGANISM: Streptococcus pneumoniae  
 US-09-815-242-13435

Query Match 2.8%; Score 6; DB 10; Length 150;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 158 AEGKAI 163  
 |||||  
 DB 63 AEGKAI 68

Search completed: April 9, 2003, 14:54:12  
 Job time : 54 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: April 9, 2003, 14:40:48 ; Search time 44 Seconds  
(without alignments)  
476.302 Million cell updates/sec

Title: US-10-034-500-2

Perfect score: 218

Sequence: 1 AEVTASCTKRVESYNYLVLDY.....QEOILVTEBVLGVNFAR 218

Scoring table: OLIGO

Searched: 283224 seqs, 96134422 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	3.7	210	2 I40272	outer surface prot
2	8	3.2	140	2 G84608	En/Spm-like transp
3	7	3.2	177	2 I40129	outer surface prot
4	7	3.2	191	2 S70262	outer surface prot
5	7	3.2	192	2 S70272	outer surface prot
6	7	3.2	196	2 S54196	outer surface prot
7	7	3.2	197	2 C82619	Heli protein Xp193
8	7	3.2	202	2 AC3318	sodium-dependent p
9	7	3.2	203	2 I40111	outer surface prot
10	7	3.2	204	2 I40112	outer surface prot
11	7	3.2	205	2 I40105	outer surface prot
12	7	3.2	209	2 S69926	outer surface prot
13	7	3.2	210	2 S69920	outer surface prot
14	7	3.2	210	2 S69925	outer surface prot
15	7	3.2	210	2 I40274	outer surface prot
16	7	3.2	210	2 I40280	outer surface prot
17	7	3.2	210	2 G70218	outer surface prot
18	7	3.2	210	2 S69927	outer surface prot
19	7	3.2	212	2 I40275	outer surface prot
20	7	3.2	213	2 B83003	glutamine amidotra
21	7	3.2	214	2 S69916	outer surface prot
22	7	3.2	223	2 D95951	probable two-compo
23	7	3.2	252	2 T35767	hypothetical prote
24	7	3.2	257	2 C75077	iron (iii) abc tra
25	7	3.2	266	2 G83178	conserved hypotet
26	7	3.2	267	2 B83858	hypothetical prote
27	7	3.2	276	2 A84309	cobalamin biosynth
28	7	3.2	304	2 F95337	probable ABC trans
29	7	3.2	311	2 A83329	probable phosphati

30	7	3.2	343	2 G64744	probable ABC-type
31	7	3.2	343	2 A90654	ATP-binding compon
32	7	3.2	343	2 A85505	ATP-binding compon
33	7	3.2	355	2 B87394	conserved hypotet
34	7	3.2	374	2 S76392	phosphoglycerate k
35	7	3.2	379	2 T04608	ADP,ATP carrier pr
36	7	3.2	407	2 D85643	hypothetical prote
37	7	3.2	422	2 A99783	probable tail tip
38	7	3.2	429	2 B64617	hypothetical prote
39	7	3.2	429	2 F71896	hypothetical prote
40	7	3.2	447	2 AD2474	hypothetical prote
41	7	3.2	482	2 C69181	proline-tRNA ligas
42	7	3.2	524	2 F84511	hypothetical prote
43	7	3.2	551	2 C84549	probable ubiquitin
44	7	3.2	552	1 A60000	nonstructural prot
45	7	3.2	554	1 A48553	glucose-6-phosphat
46	7	3.2	554	1 NDBY	probable transpor
47	7	3.2	561	2 G90744	probable transpor
48	7	3.2	561	2 C85595	probable membrane
49	7	3.2	561	2 G64822	flagellar hook-ass
50	7	3.2	634	2 F97172	membrane-associate
51	7	3.2	660	2 T04893	hypothetical prote
52	7	3.2	699	2 T18426	hypothetical prote
53	7	3.2	739	2 C96805	hypothetical prote
54	7	3.2	837	2 JN0292	antigen 332 - mala
55	7	3.2	866	2 AF1350	endopeptidase clp
56	7	3.2	1011	2 E85054	probable transpos
57	7	3.2	1071	2 T01474	hypothetical prote
58	7	3.2	1077	2 F72233	conserved hypotet
59	7	3.2	1206	2 T30567	ATP-binding caset
60	7	3.2	1426	2 S32373	DNA-binding protei
61	7	3.2	1490	2 AB2276	two-component hybr
62	7	3.2	1550	2 T24098	hypothetical prote
63	7	3.2	1603	2 AG2335	hypothetical prote
64	7	3.2	1821	2 S33441	EF protein - Strept
65	7	3.2	1822	2 A36527	juvenile-hormone e
66	6	2.8	15	2 B87458	hypothetical prote
67	6	2.8	39	2 S48690	ubiquinol-cytochro
68	6	2.8	69	1 T41731	probable proteinas
69	6	2.8	77	2 AE1645	holin from bacteri
70	6	2.8	86	2 S03117	class II histocomp
71	6	2.8	87	2 B90584	30S ribosomal prot
72	6	2.8	88	2 A84450	hypothetical prote
73	6	2.8	89	2 S52275	hypothetical prote
74	6	2.8	91	2 AE1163	hypothetical prote
75	6	2.8	93	2 AE1522	hypothetical prote
76	6	2.8	94	2 T49951	hypothetical prote
77	6	2.8	100	2 A70424	conserved hypotet
78	6	2.8	101	2 B82849	hypothetical prote
79	6	2.8	104	2 F84232	hypothetical prote
80	6	2.8	106	1 GDRB	glutaredoxin - rab
81	6	2.8	107	2 B53974	hypothetical prote
82	6	2.8	108	2 B55124	thioredoxin - Chlo
83	6	2.8	111	2 S71015	alar protein - Str
84	6	2.8	111	2 A70945	hypothetical prote
85	6	2.8	113	2 S57410	hypothetical prote
86	6	2.8	115	2 AD0872	is heavy chain V-J
87	6	2.8	118	2 C82887	conserved hypotet
88	6	2.8	118	2 D89818	purine nucleoside
89	6	2.8	121	2 A46561	7,8-dihydropter
90	6	2.8	121	2 A46561	MHC class II histo
91	6	2.8	123	2 A83041	conserved hypotet
92	6	2.8	123	2 A97138	uncharacterized pr
93	6	2.8	124	1 FAD02	profilin II - slim
94	6	2.8	126	2 UC2283	ribosomal protein
95	6	2.8	129	2 AC0071	probable membrane
96	6	2.8	130	2 AC2823	hypothetical prote
97	6	2.8	131	2 G97196	cheY-like receiver
98	6	2.8	132	1 D69934	cell wall enzyme h
99	6	2.8	132	2 D81381	transcription term
100	6	2.8	134	2 B60497	MHC class II histo
101	6	2.8	135	2 H65073	hypothetical prote
102	6	2.8	135	2 H91099	hypothetical prote

probable ABC-type  
ATP-binding compon  
ATP-binding compon  
conserved hypotet  
phosphoglycerate k  
ADP,ATP carrier pr  
hypothetical prote  
probable tail tip  
hypothetical prote  
hypothetical prote  
hypothetical prote  
proline-tRNA ligas  
hypothetical prote  
probable ubiquitin  
nonstructural prot  
glucose-6-phosphat  
probable transpor  
probable membrane  
flagellar hook-ass  
membrane-associate  
hypothetical prote  
hypothetical prote  
antigen 332 - mala  
endopeptidase clp  
probable transpos  
hypothetical prote  
conserved hypotet  
ATP-binding caset  
DNA-binding protei  
two-component hybr  
hypothetical prote  
hypothetical prote  
EF protein - Strept  
juvenile-hormone e  
hypothetical prote  
ubiquinol-cytochro  
probable proteinas  
holin from bacteri  
class II histocomp  
30S ribosomal prot  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
conserved hypotet  
hypothetical prote  
glutaredoxin - rab  
hypothetical prote  
thioredoxin - Chlo  
alar protein - Str  
hypothetical prote  
is heavy chain V-J  
conserved hypotet  
purine nucleoside  
7,8-dihydropter  
MHC class II histo  
conserved hypotet  
uncharacterized pr  
profilin II - slim  
ribosomal protein  
probable membrane  
hypothetical prote  
cheY-like receiver  
cell wall enzyme h  
transcription term  
MHC class II histo  
hypothetical prote  
hypothetical prote



103	6	2.8	135	2	D85945	hypothetical prote
104	6	2.8	139	1	F64502	hypothetical prote
105	6	2.8	140	2	AP0218	flagellar protein
106	6	2.8	141	2	I64159	hypothetical prote
107	6	2.8	141	2	S50650	hypothetical prote
108	6	2.8	141	2	D83834	flagellar biosynth
109	6	2.8	146	2	AD0524	probable PTS syste
110	6	2.8	147	2	A54492	hemoglobin beta-II
111	6	2.8	147	2	A58345	myoglobin - tremet
112	6	2.8	147	2	G87278	PTS system, IIA co
113	6	2.8	148	2	G70890	probable repressor
114	6	2.8	150	2	F95257	ribosomal protein
115	6	2.8	150	2	F98122	50S ribosomal prot
116	6	2.8	150	2	T44952	flagella-related p
117	6	2.8	152	2	AF2282	hypothetical prote
118	6	2.8	152	2	S44740	CO2C2.2 protein -
119	6	2.8	153	1	GEYL2	leghemoglobin II -
120	6	2.8	154	2	GPYL	conserved hypotet
121	6	2.8	154	2	D83516	hypothetical prote
122	6	2.8	154	2	T01820	hypothetical prote
123	6	2.8	156	2	F87255	conserved hypotet
124	6	2.8	157	2	T05720	ribosomal protein
125	6	2.8	164	2	S75589	transposase s1125
126	6	2.8	164	2	S75211	transposase s1193
127	6	2.8	170	2	D75554	hypothetical prote
128	6	2.8	171	2	S74486	hypothetical prote
129	6	2.8	171	2	T00915	hypothetical prote
130	6	2.8	171	2	G75213	hypothetical prote
131	6	2.8	175	2	I6591	interleukin 10 - p
132	6	2.8	175	2	B69848	conserved hypotet
133	6	2.8	178	2	E70305	molbdennu cofacto
134	6	2.8	178	2	T07189	late embryogenesis
135	6	2.8	178	2	H80068	hypothetical prote
136	6	2.8	179	2	C69727	biosynthesis of te
137	6	2.8	179	2	B86226	hypothetical prote
138	6	2.8	181	2	T08670	cell division cont
139	6	2.8	184	2	A36862	photosynthetic res
140	6	2.8	184	2	S41451	rega protein - Rho
141	6	2.8	186	2	T03992	hypothetical prote
142	6	2.8	186	2	T02035	hypothetical prote
143	6	2.8	188	2	A57193	oligopeptide ABC-t
144	6	2.8	189	2	G84451	probable cold-regu
145	6	2.8	189	2	S70278	outer surface prot
146	6	2.8	191	2	T24277	hypothetical prote
147	6	2.8	192	2	T21210	hypothetical prote
148	6	2.8	192	2	T46696	hypothetical prote
149	6	2.8	193	2	B96723	hypothetical prote
150	6	2.8	202	1	TWMSHS	fibroblast growth
151	6	2.8	202	2	G71022	hypothetical prote
152	6	2.8	204	2	T50736	4-vinyl protochlor
153	6	2.8	206	2	JH0145	hypothetical 24.1k
154	6	2.8	207	2	AH0219	flagellar basal-bo
155	6	2.8	209	2	I40273	outer surface prot
156	6	2.8	210	2	B84943	DNA-(apurinic or a
157	6	2.8	210	2	B71527	probable o-sialogl
158	6	2.8	211	2	I40277	outer surface prot
159	6	2.8	211	2	H90272	hypothetical prote
160	6	2.8	212	2	S70254	outer surface prot
161	6	2.8	212	2	S70254	outer surface prot
162	6	2.8	212	2	AH0643	probable lipoprote
163	6	2.8	213	2	C71088	hypothetical prote
164	6	2.8	213	2	G85673	hypothetical prote
165	6	2.8	213	2	G85673	hypothetical prote
166	6	2.8	213	2	G85673	hypothetical prote
167	6	2.8	213	2	G85673	hypothetical prote
168	6	2.8	213	2	G85673	hypothetical prote
169	6	2.8	216	2	F29504	alkylmercury lyase
170	6	2.8	216	2	B84450	hypothetical prote
171	6	2.8	218	2	T43227	alkylmercury lyase
172	6	2.8	218	2	T44508	mezB protein limp
173	6	2.8	221	2	AB3018	thiamin-phosphate
174	6	2.8	223	2	H83808	O-methyltransferas
175	6	2.8	223	2	C86827	cation transport A
176	6	2.8	225	1	S73560	phosphate transport
177	6	2.8	225	2	T05334	endomembrane-assoc
178	6	2.8	226	2	B75270	ABC transporter, A
179	6	2.8	227	2	T47416	glutathione transf
180	6	2.8	230	2	B87578	hypothetical prote
181	6	2.8	231	2	B82820	two-component syst
182	6	2.8	231	2	AB1202	E. coli copper hom
183	6	2.8	231	2	AB1159	E. coli copper hom
184	6	2.8	233	1	FVVFBFA	flavin precursor -
185	6	2.8	234	1	AG1167	hypothetical prote
186	6	2.8	238	2	F69352	branched-chain ami
187	6	2.8	238	2	F98266	thiamin-phosphate
188	6	2.8	239	2	T03783	caffeoyl-CoA O-met
189	6	2.8	239	2	B69946	phage-related term
190	6	2.8	240	2	T04084	caffeoyl-CoA O-met
191	6	2.8	242	2	T03798	caffeoyl-CoA O-met
192	6	2.8	242	2	T03798	caffeoyl-CoA O-met
193	6	2.8	242	2	T03801	caffeoyl-CoA O-met
194	6	2.8	244	2	D97281	pseudouridylylate sy
195	6	2.8	244	2	F69260	nitrate ABC transp
196	6	2.8	245	2	F71963	fumarate reductase
197	6	2.8	245	2	G64543	fumarate reductase
198	6	2.8	246	2	E83249	probable short-cha
199	6	2.8	246	2	C81946	hypothetical prote
200	6	2.8	247	2	JT0555	MHC class II histo
201	6	2.8	247	2	T09757	caffeoyl-CoA O-met
202	6	2.8	247	2	T02920	caffeoyl-CoA O-met
203	6	2.8	247	2	T09399	caffeoyl-CoA O-met
204	6	2.8	247	2	AB1957	1-acylglycerol-3-p
205	6	2.8	247	2	G81013	1-acyl-sn-glycerol
206	6	2.8	249	2	S45439	cytochrome-c oxida
207	6	2.8	249	2	C84270	30S ribosomal prot
208	6	2.8	249	2	JQ1981	lectin II - Scotch
209	6	2.8	249	2	T10731	caffeoyl-CoA O-met
210	6	2.8	250	2	A70678	hypothetical prote
211	6	2.8	250	2	A69725	triase-phosphate i
212	6	2.8	253	2	H89845	hypothetical prote
213	6	2.8	254	2	T12206	caffeoyl-CoA O-met
214	6	2.8	254	2	T22677	hypothetical prote
215	6	2.8	254	2	D64540	signal-transducing
216	6	2.8	258	2	T11582	lipopeptide antibi
217	6	2.8	259	2	T05431	probable caffeoyl-
218	6	2.8	260	2	G75111	hypothetical prote
219	6	2.8	260	2	G75114	hypothetical prote
220	6	2.8	260	2	G87349	conserved hypotet
221	6	2.8	261	2	E95963	probable transcrip
222	6	2.8	261	2	C75058	hypothetical prote
223	6	2.8	263	2	HLMSER	H-2 class II histo
224	6	2.8	264	1	HLMSER	H-2 class II histo
225	6	2.8	264	1	HLMSER	H-2 class II histo
226	6	2.8	264	2	A60497	H-2 class II histo
227	6	2.8	264	2	S10989	class II histocomp
228	6	2.8	264	2	S11650	class II histocomp
229	6	2.8	264	2	A50529	H-2 class II histo
230	6	2.8	264	2	I48422	MHC E-beta-f - mou
231	6	2.8	264	2	I56056	MHC E-beta-f - mou
232	6	2.8	264	2	F83351	conserved hypotet
233	6	2.8	264	2	C95092	PEP-utilizing enzy
234	6	2.8	265	2	C71032	hypothetical prote
235	6	2.8	265	2	D75181	hypothetical prote
236	6	2.8	267	2	JC4857	hepatocarcinogenes
237	6	2.8	271	2	S48426	RH2 protein - yea
238	6	2.8	271	2	AE2383	hypothetical prote
239	6	2.8	272	2	F70979	hypothetical prote
240	6	2.8	272	2	A72512	probable hypoxanth
241	6	2.8	273	2	AH0030	shikimate 5-dehydr
242	6	2.8	274	2	F86624	outer membrane pro
243	6	2.8	274	2	C72001	probable outer mem
244	6	2.8	278	2	T47992	hypothetical prote
245	6	2.8	278	2	T34461	hypothetical prote
246	6	2.8	280	1	A71032	hypothetical prote
247	6	2.8	281	2	C83635	hypothetical prote
248	6	2.8	282	2	C64435	shikimate 5-dehydr

249	6	2.8	282	2	B75131	3-methyl-2-oxobuta	322	6	2.8	321	2	A37842	hypothetical prote
250	6	2.8	282	2	S76306	transposase s11043	323	6	2.8	322	2	F75365	conserved hypochet
251	6	2.8	282	2	S76382	transposase s1023	324	6	2.8	322	2	AG1923	hypothetical prote
252	6	2.8	282	2	S76312	transposase s1035	325	6	2.8	323	2	T52563	probable DNA-(apur
253	6	2.8	282	2	S77237	transposase s11135	326	6	2.8	325	2	S57089	hypothetical prote
254	6	2.8	282	2	S75484	transposase s11199	327	6	2.8	325	2	AE1134	oxetanocin A resis
255	6	2.8	282	2	F82748	hypothetical prote	328	6	2.8	327	2	F82748	hypothetical prote
256	6	2.8	282	2	H82748	hypothetical prote	329	6	2.8	327	2	F95889	probable dehydroge
257	6	2.8	282	2	H75250	OxR1-related prote	330	6	2.8	328	2	E75514	B-cell receptor as
258	6	2.8	282	2	F96946	methyl-accepting c	331	6	2.8	329	1	VMU21	VSG expression sit
259	6	2.8	283	2	B71086	probable 3-methyl-	332	6	2.8	329	2	AC3221	oxidoreductase Atu
260	6	2.8	283	2	C82749	hypothetical prote	333	6	2.8	330	2	G64345	integrase homolog
261	6	2.8	283	2	D82749	hypothetical prote	334	6	2.8	332	2	E95033	PTS system, mann
262	6	2.8	283	2	T36962	hypothetical prote	335	6	2.8	332	2	E97904	phosphotransfere
263	6	2.8	285	2	AE1684	panthoemate synth	336	6	2.8	333	2	C84620	hypothetical prote
264	6	2.8	286	2	B83343	hypothetical prote	337	6	2.8	333	2	F85035	hypothetical prote
265	6	2.8	286	2	E69465	conserved hypochet	338	6	2.8	334	2	A71480	probable glycerol-
266	6	2.8	288	2	T21770	hypothetical prote	339	6	2.8	334	2	D72381	dehydrogenase - Th
267	6	2.8	289	2	T34688	probable lipoprote	340	6	2.8	334	2	E96945	sensory transducti
268	6	2.8	290	2	A61397	N-hydroxyarylamine	341	6	2.8	334	2	AD1049	HLIC protein (EC 3
269	6	2.8	290	2	I67465	arylamine N-acetyl	342	6	2.8	334	2	G86113	probable integral
270	6	2.8	290	2	I58425	arylamine N-acetyl	343	6	2.8	334	2	G91272	probable integral
271	6	2.8	290	2	S66357	mannose/glucose-bi	344	6	2.8	334	2	C43653	probable integral
272	6	2.8	290	2	AP3234	phenazine biosynth	345	6	2.8	335	2	T23050	hypothetical prote
273	6	2.8	291	2	A86681	hypothetical prote	346	6	2.8	335	2	T35941	hypothetical prote
274	6	2.8	291	2	H97959	hypothetical prote	347	6	2.8	336	2	B90209	conserved hypochet
275	6	2.8	293	2	B90865	hypothetical prote	348	6	2.8	337	1	H70121	methanol dehydroge
276	6	2.8	293	2	G85753	hypothetical prote	349	6	2.8	337	2	AH3207	transcription regu
277	6	2.8	293	2	B64880	probable multiple	350	6	2.8	338	2	S77605	glutamate/glutamin
278	6	2.8	294	2	S60991	hypothetical prote	351	6	2.8	338	2	T24933	hypothetical prote
279	6	2.8	294	2	B36351	sarco toxin II-2 -	352	6	2.8	341	2	E96941	ADP-ribosylglycohy
280	6	2.8	294	2	C86802	hypothetical prote	353	6	2.8	342	2	T21098	DNA (apurinic or a
281	6	2.8	297	2	E97350	ABC-type sugar tra	354	6	2.8	342	2	T41498	conserved hypochet
282	6	2.8	297	2	B84237	hypothetical prote	355	6	2.8	342	2	C75172	ATP-dependent prote
283	6	2.8	297	2	A89451	protein T0469.4 [I	356	6	2.8	342	2	E87605	hypothetical prote
284	6	2.8	298	2	A69623	flagellar hook-ass	357	6	2.8	342	2	T26792	ATP-dependent prote
285	6	2.8	298	2	S53483	probable membrane	358	6	2.8	343	2	A99353	hypothetical prote
286	6	2.8	298	2	B71013	hypothetical prote	359	6	2.8	343	2	T48020	transcription acti
287	6	2.8	300	2	D85431	hypothetical prote	360	6	2.8	344	2	JS0571	lima protein - Pse
288	6	2.8	302	2	C64829	transcription regu	361	6	2.8	344	2	B39133	hypothetical prote
289	6	2.8	303	2	AD1367	Erwinia chrysanth	362	6	2.8	344	2	G75057	phosphoserine phos
290	6	2.8	303	2	AE1736	Erwinia chrysanth	363	6	2.8	344	2	B69517	anthranilate phosp
291	6	2.8	304	2	T18345	MGC2 protein precu	364	6	2.8	345	2	F90239	conjugal transfer
292	6	2.8	305	2	T07049	probable epoxide h	365	6	2.8	347	2	G82862	probable oxidoredu
293	6	2.8	307	2	H70720	hypothetical prote	366	6	2.8	347	2	A95983	protein plex - Syn
294	6	2.8	307	2	T27212	hypothetical prote	367	6	2.8	348	1	S75456	DNA (apurinic or a
295	6	2.8	307	2	C71002	hypothetical prote	368	6	2.8	349	2	T41497	gcpE protein [limp
296	6	2.8	308	2	T17524	hypothetical prote	369	6	2.8	349	2	G97121	hypothetical prote
297	6	2.8	308	2	D83858	hypothetical prote	370	6	2.8	350	2	T21106	hypothetical prote
298	6	2.8	309	2	T41203	WD repeat protein	371	6	2.8	351	2	JQ2327	AlI protein - Indi
299	6	2.8	310	2	T44857	probable hydroxyla	372	6	2.8	351	2	T43111	hypothetical prote
300	6	2.8	310	2	T00740	DNA-directed DNA p	373	6	2.8	353	2	H97636	N5,N10-methylene
301	6	2.8	310	2	F82979	probable transcrip	374	6	2.8	353	2	AB2860	monooxygenase [limp
302	6	2.8	310	2	T36542	hypothetical prote	375	6	2.8	356	2	A99901	D-alanine-D-alanin
303	6	2.8	310	2	I52646	DNA binding protei	376	6	2.8	357	1	QQCVCI	AlI protein - toma
304	6	2.8	312	2	F71922	acetyl-coenzyme A	377	6	2.8	357	2	S46795	hypothetical prote
305	6	2.8	312	2	E64589	acetyl-CoA carboxy	378	6	2.8	357	2	A12938	hypothetical prote
306	6	2.8	312	2	C83077	riboflavin kinase/	379	6	2.8	358	2	S07594	hypothetical prote
307	6	2.8	313	2	AD1352	S. aureus CbfI pro	380	6	2.8	358	2	B86141	protein T25K16.3 [
308	6	2.8	313	2	AG1722	S. aureus CbfI pro	381	6	2.8	359	2	T34602	ATP-dependent RNA
309	6	2.8	314	2	D70313	riboflavin kinase	382	6	2.8	360	2	G90370	probable 2-isoprop
310	6	2.8	314	2	AH2551	chromosome particl	383	6	2.8	361	2	B71181	hypothetical prote
311	6	2.8	316	2	T46282	hypothetical prote	384	6	2.8	362	2	AB4187	hypothetical prote
312	6	2.8	317	2	S57275	triacylglycerol li	385	6	2.8	363	2	B86248	protein T23J18.11
313	6	2.8	317	2	F45774	odorant receptor 4	386	6	2.8	364	2	AF2459	hypothetical prote
314	6	2.8	319	1	OWBS	ornithine carbamoy	387	6	2.8	365	2	S10847	collagen alpha 2(I
315	6	2.8	319	2	D64303	hypothetical prote	388	6	2.8	365	2	T23220	hypothetical prote
316	6	2.8	320	2	T43040	hypothetical prote	389	6	2.8	366	2	S46305	dihydrodipicolinat
317	6	2.8	320	1	GERTS	bone sialoprotein	390	6	2.8	366	2	D75167	hypothetical prote
318	6	2.8	321	2	T38413	probable oxidoredu	391	6	2.8	367	2	G95197	alanine racemase [
319	6	2.8	321	2	T07048	probable epoxide h	392	6	2.8	367	2	F98343	miIX protein (AF00
320	6	2.8	321	2	T07043	probable epoxide h	393	6	2.8	367	2	A83825	peptidoglycan acet
321	6	2.8	321	2	T07044	probable epoxide h	394	6	2.8	368	2	AD3493	leucine-, isoleuci

395	6	2.8	370	2	D87692	468	6	2.8	424	2	S74788	hypothetical prote
396	6	2.8	370	2	AC1272	469	6	2.8	425	2	G86217	protein T2767.14 l
397	6	2.8	370	2	AE1634	470	6	2.8	426	2	A81933	probable malate de
398	6	2.8	370	2	G87621	471	6	2.8	426	2	C81173	malate oxidoreduc
399	6	2.8	370	2	T31533	472	6	2.8	427	2	A64329	phosphorylase hy
400	6	2.8	370	2	B97671	473	6	2.8	427	2	T01505	hypothetical prote
401	6	2.8	370	2	AG2895	474	6	2.8	427	2	AH2350	hypothetical prote
402	6	2.8	371	2	T36876	475	6	2.8	428	2	T46046	hypothetical prote
403	6	2.8	371	2	S55588	476	6	2.8	429	2	AB3350	NAD(P) transhydro
404	6	2.8	372	2	B34251	477	6	2.8	429	2	A81315	probable MCP-type
405	6	2.8	372	2	C98064	478	6	2.8	431	2	H82337	histidinol dehydro
406	6	2.8	375	2	AG0350	479	6	2.8	434	2	B72410	alkaline phosphat
407	6	2.8	375	2	C95396	480	6	2.8	434	2	T43197	nucleosporin homo
408	6	2.8	377	2	B95944	481	6	2.8	437	2	AG0111	X-Pro antipeptida
409	6	2.8	377	2	C69955	482	6	2.8	439	2	T31124	hypothetical prote
410	6	2.8	377	2	C90296	483	6	2.8	441	2	T35083	histidinol dehydro
411	6	2.8	378	2	G85342	484	6	2.8	441	2	T42946	hypothetical prote
412	6	2.8	378	2	G97301	485	6	2.8	442	2	C96672	hypothetical prote
413	6	2.8	379	2	B97147	486	6	2.8	444	2	AC2755	GTP-binding protei
414	6	2.8	380	2	AH3647	487	6	2.8	445	2	S56207	EMP47 protein prec
415	6	2.8	380	2	H90261	488	6	2.8	446	2	T09977	H+-transporting tw
416	6	2.8	382	2	A88089	489	6	2.8	447	2	S32227	glutamate dehydro
417	6	2.8	382	2	T34127	490	6	2.8	447	2	C82958	probable two-compo
418	6	2.8	384	2	AE1987	491	6	2.8	447	2	A97211	glycosyltransferas
419	6	2.8	384	2	G96738	492	6	2.8	451	2	T23731	hypothetical prote
420	6	2.8	385	2	AH3601	493	6	2.8	452	2	A75523	cldd protein PA801
421	6	2.8	385	2	S64614	494	6	2.8	452	2	B90407	conserved hypochet
422	6	2.8	385	2	B83981	495	6	2.8	452	2	S46037	prephenate dehydro
423	6	2.8	387	2	B83981	496	6	2.8	454	2	S18625	nitr protein - Azo
424	6	2.8	387	2	F70886	497	6	2.8	454	2	D75446	oxidoreductase - p
425	6	2.8	387	2	F70886	498	6	2.8	455	2	F97065	D-xylulose-proton sy
426	6	2.8	391	2	E87258	499	6	2.8	455	2	S61158	hypothetical prote
427	6	2.8	392	2	B23270	500	6	2.8	456	2	T09290	late embryonic abu
428	6	2.8	392	2	B97006	501	6	2.8	457	2	T18673	hypothetical prote
429	6	2.8	394	1	K1BSGM	502	6	2.8	458	2	S75328	DNA-directed DNA p
430	6	2.8	394	2	C69675	503	6	2.8	461	2	T08041	phosphoglycerate k
431	6	2.8	394	2	JO1399	504	6	2.8	465	2	S68987	transcription acti
432	6	2.8	395	2	C71148	505	6	2.8	468	2	A49131	inositol 1,4,5-tri
433	6	2.8	396	2	A83529	506	6	2.8	468	2	C96818	hypothetical prote
434	6	2.8	399	2	T10002	507	6	2.8	468	2	AE1942	two-component sens
435	6	2.8	399	2	T12126	508	6	2.8	470	2	AF1972	hypothetical prote
436	6	2.8	403	2	H81247	509	6	2.8	472	2	S39719	spore coat polysac
437	6	2.8	403	2	T39846	510	6	2.8	473	2	H72744	probable D-lactate
438	6	2.8	403	2	T38078	511	6	2.8	475	1	JN0327	sulfate adenylyl tr
439	6	2.8	403	2	T49480	512	6	2.8	475	2	E91079	hypothetical prote
440	6	2.8	405	2	H69711	513	6	2.8	475	2	F85924	hypothetical prote
441	6	2.8	407	2	AH2594	514	6	2.8	475	2	T38597	zinc finger protei
442	6	2.8	407	2	H72573	515	6	2.8	476	2	G82062	sulfate adenylylase
443	6	2.8	407	2	B71679	516	6	2.8	476	2	T40086	hypothetical prote
444	6	2.8	407	2	F70764	517	6	2.8	476	2	AH1044	aspartate ammonia-
445	6	2.8	408	2	S63528	518	6	2.8	478	2	AG0408	sulfate adenylyl tr
446	6	2.8	408	2	P90356	519	6	2.8	478	2	T32476	probable protein-t
447	6	2.8	408	2	B97741	520	6	2.8	479	2	AT0856	AMP sulfurylase (A
448	6	2.8	409	2	B87262	521	6	2.8	481	2	AH3350	alkaline phosphat
449	6	2.8	409	2	H70524	522	6	2.8	486	2	T36346	probable aldehyde
450	6	2.8	410	1	KHMSD	523	6	2.8	486	2	A56145	cardiolipin synthe
451	6	2.8	412	2	H75206	524	6	2.8	486	2	AG0266	cardiolipin synthe
452	6	2.8	413	2	H70606	525	6	2.8	486	2	D85705	cardiolipin synthe
453	6	2.8	413	2	T35502	526	6	2.8	486	2	P90847	cardiolipin synthe
454	6	2.8	414	2	T11617	527	6	2.8	486	2	AF0651	cardiolipin synthe
455	6	2.8	414	2	H87191	528	6	2.8	486	2	C83971	UDP-N-acetylmuramo
456	6	2.8	414	2	D96838	529	6	2.8	489	2	S66564	acetyl CoA carboxy
457	6	2.8	415	2	H90198	530	6	2.8	489	2	AC2343	hypothetical prote
458	6	2.8	415	2	H82095	531	6	2.8	489	2	S50396	hypothetical prote
459	6	2.8	416	2	A85112	532	6	2.8	492	2	P86384	probable protein f
460	6	2.8	417	1	S20608	533	6	2.8	492	2	AB0883	probable aldehyde
461	6	2.8	417	1	A40968	534	6	2.8	497	2	T52308	very-long-chain fa
462	6	2.8	417	1	A42843	535	6	2.8	498	2	B75373	GsDEF family prote
463	6	2.8	417	2	I52968	536	6	2.8	501	2	JC7181	maternal transcrip
464	6	2.8	418	2	E36687	537	6	2.8	502	2	C83857	anthranilate synth
465	6	2.8	420	2	E75274	538	6	2.8	502	2	P87436	peptidoglycan bind
466	6	2.8	422	2	S52790	539	6	2.8	505	2	B83918	carboxypeptidase B
467	6	2.8	422	2	S52790	540	6	2.8	506	2	AD1119	flavocytochrome C

541	6	2.8	506	2	AG1479	614	6	2.8	598	2	A96685	probable RING zinc
542	6	2.8	506	2	T05272	615	6	2.8	598	2	T38403	probable nucleopor
543	6	2.8	508	2	AB1147	616	6	2.8	599	2	S15004	acetylactate synth
544	6	2.8	508	2	BB1875	617	6	2.8	600	2	S48509	transcription regu
545	6	2.8	508	2	T45867	618	6	2.8	603	2	AB1967	two-component hybr
546	6	2.8	508	2	A95903	619	6	2.8	603	2	AD2397	ATP-binding protei
547	6	2.8	509	2	T45868	620	6	2.8	610	2	F95898	probable nodulari
548	6	2.8	512	2	B64119	621	6	2.8	611	2	G90542	gfp-binding protei
549	6	2.8	512	2	D82444	622	6	2.8	611	2	T19217	hypothetical prote
550	6	2.8	512	2	P84538	623	6	2.8	611	2	T29487	hypothetical prote
551	6	2.8	513	2	T38770	624	6	2.8	613	2	H90160	conserved hypotnet
552	6	2.8	514	2	A70685	625	6	2.8	615	1	S06977	nif-specific regul
553	6	2.8	517	2	A13201	626	6	2.8	615	2	AH2028	hypothetical prote
554	6	2.8	518	2	AF0937	627	6	2.8	621	2	D72491	hypothetical prote
555	6	2.8	520	2	F69470	628	6	2.8	622	2	B69609	probable heavy-met
556	6	2.8	521	2	T16692	629	6	2.8	622	2	H96684	cytochrome-c oxid
557	6	2.8	522	2	T18583	630	6	2.8	622	2	T40742	probable RING zinc
558	6	2.8	523	2	T20828	631	6	2.8	625	2	T40742	hypothetical integ
559	6	2.8	523	2	B70211	632	6	2.8	628	2	T33792	hypothetical prote
560	6	2.8	530	2	AF2175	633	6	2.8	630	2	H83200	probable acetyl-co
561	6	2.8	531	2	T20763	634	6	2.8	632	2	D82213	DNA endonuclease
562	6	2.8	531	2	T45865	635	6	2.8	632	2	T22288	pppe protein VC134
563	6	2.8	531	2	T49058	636	6	2.8	632	2	T22288	hypothetical prote
564	6	2.8	532	2	QEPGM	637	6	2.8	632	2	B84157	glucose-inhibited
565	6	2.8	532	2	B83813	638	6	2.8	637	2	D83091	ATP sulfurylase GT
566	6	2.8	533	2	B83813	639	6	2.8	638	2	A36929	hypothetical prote
567	6	2.8	536	2	B75329	640	6	2.8	640	2	D83571	virulence regulato
568	6	2.8	538	2	AB1856	641	6	2.8	641	2	AG2164	conserved hypotnet
569	6	2.8	541	2	T35143	642	6	2.8	642	2	F72528	hypothetical prote
570	6	2.8	541	2	T45563	643	6	2.8	643	2	S17997	probable Glu-tRNA
571	6	2.8	542	2	S76617	644	6	2.8	644	2	S55395	hypothetical gene
572	6	2.8	545	2	A49229	645	6	2.8	646	2	AE2428	nucleofilament prot
573	6	2.8	546	2	S55386	646	6	2.8	646	2	F90111	hypothetical prote
574	6	2.8	547	2	B69647	647	6	2.8	651	2	I37224	DNA primase (import
575	6	2.8	548	2	T36881	648	6	2.8	653	2	S17691	Rab geranylgeranyl
576	6	2.8	548	2	T41182	649	6	2.8	655	2	C71438	acetylactate synth
577	6	2.8	551	2	B69371	650	6	2.8	657	2	S32991	hypothetical prote
578	6	2.8	551	2	B90247	651	6	2.8	657	2	G71377	hypothetical prote
579	6	2.8	552	1	NMXR81	652	6	2.8	660	2	T45588	probable transcrip
580	6	2.8	552	1	A44277	653	6	2.8	662	2	E64650	arm repeat contain
581	6	2.8	552	1	B44277	654	6	2.8	662	2	AE3290	acetyl-CoA synthet
582	6	2.8	552	1	C44277	654	6	2.8	663	2	AE3290	propionate-CoA lig
583	6	2.8	552	1	B60000	655	6	2.8	663	2	I46611	Mx protein - pig
584	6	2.8	552	1	S06895	656	6	2.8	664	2	H70396	cation transportin
585	6	2.8	553	2	S53080	657	6	2.8	665	2	H87468	ubiquinol oxidase
586	6	2.8	554	2	F86244	658	6	2.8	666	2	F58931	NMDH2 dehydrogenas
587	6	2.8	555	2	T38108	659	6	2.8	666	2	G82130	2,4-dienoyl-CoA re
588	6	2.8	555	2	AG1914	660	6	2.8	666	2	G90566	hypothetical prote
589	6	2.8	556	2	H84125	661	6	2.8	667	2	B86843	hypothetical prote
590	6	2.8	556	2	G88941	662	6	2.8	669	2	D96006	probable cytochrom
591	6	2.8	557	1	DWPSUP	663	6	2.8	671	2	D72346	chemotaxis sensor
592	6	2.8	557	2	T45008	664	6	2.8	672	2	C72404	glycine-tRNA ligas
593	6	2.8	557	2	T43657	665	6	2.8	674	2	B71438	hypothetical prote
594	6	2.8	559	2	B83007	666	6	2.8	677	2	S65573	phosphate-specific
595	6	2.8	560	2	AS4155	667	6	2.8	678	2	A70762	probable PPS prote
596	6	2.8	562	2	SS4464	668	6	2.8	685	2	S65974	conserved hypotnet
597	6	2.8	562	2	D89789	669	6	2.8	686	2	T23721	hypothetical prote
598	6	2.8	570	2	T24152	670	6	2.8	687	1	FYFPM	white protein - fr
599	6	2.8	571	2	C84038	671	6	2.8	688	2	D84248	hypothetical prote
600	6	2.8	571	2	C84356	672	6	2.8	688	2	H70024	sorbitol-6-phospha
601	6	2.8	573	2	B70169	673	6	2.8	689	2	F83843	hypothetical prote
602	6	2.8	580	2	C71182	674	6	2.8	690	2	B86442	probable PPR-repea
603	6	2.8	582	2	A90192	675	6	2.8	695	2	S37439	tranketolase (BC
604	6	2.8	585	2	B88087	676	6	2.8	695	2	G87316	nuclease, probable
605	6	2.8	588	2	E75060	677	6	2.8	701	2	A11501	internalin protein
606	6	2.8	588	2	T45564	678	6	2.8	706	2	G90696	hypothetical prote
607	6	2.8	588	2	T45539	679	6	2.8	708	2	A47176	probable transmemb
608	6	2.8	589	2	S60154	680	6	2.8	709	2	C28821	1-phosphatidylinos
609	6	2.8	589	2	S60159	681	6	2.8	713	1	S30019	acetate-CoA ligase
610	6	2.8	589	2	T19216	682	6	2.8	717	1	XNBYGM	glutamine-fructose
611	6	2.8	589	2	D95953	683	6	2.8	720	2	C85547	probable cytoplas
612	6	2.8	597	2	A95949	684	6	2.8	722	2	C71411	hypothetical prote
613	6	2.8	597	2	T35746	685	6	2.8	723	2	T49985	oligopeptidase A-1
						686	6	2.8	727	2	T00383	KIAA0632 protein -

687	2.8	731	2	C84464	Mutator-like trans	760	2.8	961	2	T03467	NADH dehydrogenase
688	2.8	734	1	MZBEAL	DNA-packaging prot	761	2.8	963	2	A53984	membrane alanyl am
689	2.8	743	1	H96713	hypothetical prote	762	2.8	963	2	AD2381	type I site-specific
690	2.8	745	1	PC4183	1-phosphatidylinos	763	2.8	964	1	CGCH2S	collagen alpha 2(I
691	2.8	745	2	B71654	polyrribonucleotide	764	2.8	967	1	SYMTPAT	alanine-cRNA ligas
692	2.8	746	1	QRECFE	ferriterococellin	765	2.8	967	2	A30325	membrane alanyl am
693	2.8	746	2	G90706	hypothetical prote	766	2.8	968	1	T29466	hypothetical prote
694	2.8	746	2	B85557	hypothetical prote	767	2.8	988	1	S35362	protein kinase C (
695	2.8	746	2	T29646	hypothetical prote	768	2.8	990	2	E86272	probable AAA-type
696	2.8	751	2	AC0574	ferrienterobactin	769	2.8	996	2	D84561	probable AAA-type
697	2.8	752	1	C2HU	complement C2 prec	770	2.8	997	2	T39521	hypothetical prote
698	2.8	756	1	A55943	1-phosphatidylinos	771	2.8	1003	2	T19638	hypothetical prote
699	2.8	756	1	B28821	1-phosphatidylinos	772	2.8	1005	2	T18537	Ig heavy chain - C
700	2.8	759	2	S77224	single-strand DNA-	773	2.8	1007	2	H81670	conserved hypotnet
701	2.8	760	1	C2MS	classical-compleme	774	2.8	1010	2	T36383	probable large ATP
702	2.8	762	2	JC7174	N,N-dimethylformam	775	2.8	1014	2	T18759	hypothetical prote
703	2.8	763	2	I50807	complement factor	776	2.8	1018	2	AH0293	conserved hypotnet
704	2.8	763	2	AD0170	probable membrane	777	2.8	1022	2	T51257	calmodulin-binding
705	2.8	764	2	T07608	chloride channel p	778	2.8	1022	2	S50928	calmodulin-binding
706	2.8	770	2	T50308	probable translati	779	2.8	1024	2	S18251	collagen alpha 1(X
707	2.8	775	2	T37837	probable signal tr	780	2.8	1028	2	I58164	BIG-1 protein - ra
708	2.8	779	1	MMVZAL	ribonucleoside-dip	781	2.8	1034	2	T30551	beta-galactosidase
709	2.8	780	2	T02939	chloride channel p	782	2.8	1034	2	T30574	beta-galactosidase
710	2.8	787	2	B70535	probable membrane	783	2.8	1036	2	T23845	hypothetical prote
711	2.8	795	2	B83608	hypothetical prote	784	2.8	1038	2	AG2187	hypothetical prote
712	2.8	797	2	T33098	hypothetical prote	785	2.8	1042	2	T29307	ATP-dependent dNDN
713	2.8	799	2	C82929	ATP synthase alpha	786	2.8	1046	2	A86790	tail tip fiber pro
714	2.8	805	2	A46266	aryl hydrocarbon r	787	2.8	1061	2	T33107	protein F41H10.4 [
715	2.8	809	2	A12747	conserved hypotnet	788	2.8	1061	2	C98680	neurofilament trip
716	2.8	810	2	S67050	probable membrane	789	2.8	1072	1	A37221	hypothetical prote
717	2.8	821	2	AB1126	internalin, peptid	790	2.8	1072	2	T18802	hypothetical prote
718	2.8	822	2	C72023	chltr probable pho	791	2.8	1077	2	S66842	hypothetical prote
719	2.8	822	2	B81518	conserved hypotnet	792	2.8	1078	2	E75407	iso-leucyl-tRNA syn
720	2.8	822	2	B86601	CHITr probable pho	793	2.8	1081	2	B81303	probable membrane
721	2.8	824	2	D70337	poly A polymerase	794	2.8	1086	1	B46335	HIV-1 retropepsin
722	2.8	825	2	US0174	cellulase (EC 3.2.	795	2.8	1087	2	U01162	Pol protein - Maed
723	2.8	827	2	S75622	hypothetical prote	796	2.8	1093	2	T50652	AP-3 complex betas
724	2.8	828	2	T22367	hypothetical prote	797	2.8	1094	2	S49313	protein kinase - s
725	2.8	836	2	E96561	hypothetical prote	798	2.8	1094	2	T50651	AP3-complex beta-3
726	2.8	836	2	AD2565	hypothetical prote	799	2.8	1115	2	T41342	probable colled-co
727	2.8	837	2	T23946	hypothetical prote	800	2.8	1121	2	G64103	exodeoxyribonuclea
728	2.8	838	2	B83150	probable ATP-depen	801	2.8	1146	2	S40311	integrin - fruit f
729	2.8	840	2	F82937	DNA topoisomerase	802	2.8	1148	2	UC5984	camp-dependent cel
730	2.8	840	2	H86429	hypothetical prote	803	2.8	1171	2	G97174	DNA-directed RNA p
731	2.8	845	2	E72652	pyruvate, water di	804	2.8	1175	2	S26874	DNA-directed RNA p
732	2.8	845	2	A45669	neurofilament trip	805	2.8	1185	2	H64526	proline dehydrogen
733	2.8	849	2	S00030	neurofilament trip	806	2.8	1185	2	B71980	proline dehydrogen
734	2.8	852	2	T01364	homeodomain trans	807	2.8	1189	2	T51491	hypothetical prote
735	2.8	852	2	B84001	ribonucleoside-dip	808	2.8	1191	2	S70963	pyruvate (flavodox
736	2.8	854	2	S02003	neurofilament trip	809	2.8	1195	2	A59016	DNA-directed DNA p
737	2.8	856	2	B81399	probable periplasm	810	2.8	1195	2	S38174	probable purine nu
738	2.8	859	2	S69700	hypothetical prote	811	2.8	1198	2	T49726	hypothetical prote
739	2.8	866	2	A11720	endopeptidase Clp	812	2.8	1213	2	E69255	mannosyltransferas
740	2.8	870	2	A96637	hypothetical prote	813	2.8	1217	2	F97177	alpha-glucosidase
741	2.8	873	2	H97528	hypothetical prote	814	2.8	1224	2	T26377	hypothetical prote
742	2.8	874	2	S25530	glycoprotein B - b	815	2.8	1249	2	T47885	beta-tubulin cofac
743	2.8	886	2	S73184	phycobilisome link	816	2.8	1252	2	H97178	ATP-dependent exon
744	2.8	886	2	A48586	suppressor of hair	817	2.8	1262	2	T22523	hypothetical prote
745	2.8	888	2	S61619	STD protein - yea	818	2.8	1273	2	C96767	unknown protein P2
746	2.8	895	1	IUXLCP	EP-cadherin precur	819	2.8	1274	2	T16251	hypothetical prote
747	2.8	901	2	C96696	protein PIN21.6 [I	820	2.8	1287	2	S55692	multitng resistan
748	2.8	905	2	T02205	lu-ECAM-1 protein	821	2.8	1307	2	T19497	hypothetical prote
749	2.8	906	2	AG1957	hypothetical prote	822	2.8	1337	1	I38670	protein-tyrosine-p
750	2.8	920	2	C70668	probable mmpL7 pro	823	2.8	1364	2	T42697	hypothetical prote
751	2.8	920	2	AD2143	two-component sens	824	2.8	1364	2	T00250	MEGF2 protein - hu
752	2.8	921	2	UC4625	inter-alpha-trypsi	825	2.8	1371	1	VCBEW7	major capaid prote
753	2.8	926	2	A54142	nucleoporin NUP107	826	2.8	1371	1	T42938	collagen alpha 2(I
754	2.8	929	2	I51027	type XII collagen	827	2.8	1373	1	A43291	DNA-directed RNA p
755	2.8	930	2	JX0368	inter-alpha-trypsi	828	2.8	1396	2	F87311	probable membrane
756	2.8	932	2	UCS953	inter-alpha-inhibi	829	2.8	1407	2	S59823	multitng resistan
757	2.8	937	2	A32879	oxoglutarate dehyd	830	2.8	1408	2	T43261	probable ppe prote
758	2.8	955	2	S46651	beta-agarase - Vib	831	2.8	1436	2	B70520	DNA polymerase III
759	2.8	959	2	H69344	hypothetical prote	832	2.8	1442	2	C82898	

833	6	2.8	1481	2	S78373	DNA-directed RNA p	506	5	2.3	38	2	D69677	phosphatase (RapG)
834	6	2.8	1505	2	S28079	DNA-directed DNA p	907	5	2.3	39	2	C84197	anthranilate synth
835	6	2.8	1535	2	S46224	peroxidasein - frui	908	5	2.3	32	2	A43591	43k outer membrane
836	6	2.8	1556	2	S76781	glutamate synthase	909	5	2.3	39	2	E69677	phosphatase (RapI)
837	6	2.8	1558	2	AB2457	two-component hybr	910	5	2.3	41	2	F82459	hypothetical prote
838	6	2.8	1572	2	S45251	SNP2alpha protein	911	5	2.3	42	2	T07077	heat shock protein
839	6	2.8	1586	2	S39560	HBW protein - hum	912	5	2.3	43	2	S31012	gene 67 protein -
840	6	2.8	1628	2	F86194	hypothetical prote	913	5	2.3	44	2	B29137	aad leader peptid
841	6	2.8	1634	2	JCS500	phosphoinositide 3	914	5	2.3	44	2	D81921	hypothetical prote
842	6	2.8	1635	2	A10452	hemolysin [limpote	915	5	2.3	44	2	S68844	1,2-beta-fructan 1
843	6	2.8	1649	2	C86832	hypothetical prote	916	5	2.3	46	2	T42016	ppl-like Ser/Thr p
844	6	2.8	1661	2	T21986	hypothetical prote	917	5	2.3	46	2	S23300	photosystem I chai
845	6	2.8	1663	2	T21993	hypothetical prote	918	5	2.3	46	2	A10173	hypothetical prote
846	6	2.8	1674	2	G96736	hypothetical prote	919	5	2.3	47	2	T01680	sepiapterin reduct
847	6	2.8	1679	2	T15968	hypothetical prote	920	5	2.3	48	2	B85935	hypothetical prote
848	6	2.8	1702	2	E86318	hypothetical prote	921	5	2.3	50	2	E90557	hypothetical prote
849	6	2.8	1720	2	T07258	cell division prot	922	5	2.3	51	1	KKPOT	proteinase inhibit
850	6	2.8	1729	2	A49282	fusion protein Ia/	923	5	2.3	51	2	T07336	hypothetical prote
851	6	2.8	1731	2	B98241	hypothetical prote	924	5	2.3	51	2	C82586	hypothetical prote
852	6	2.8	1731	2	AB3045	ice nucleation pro	925	5	2.3	52	2	G83717	hypothetical prote
853	6	2.8	1792	2	T08878	superfyllin P205 -	926	5	2.3	53	2	C30606	T-cell receptor al
854	6	2.8	1802	2	T00020	bacterial blight-r	927	5	2.3	54	2	B56657	pEMp2/MSA (clone
855	6	2.8	1826	2	D72120	excinuclease ABC,	928	5	2.3	54	2	S32360	SNAP receptor - bo
856	6	2.8	1938	2	JCS421	smooth muscle myos	929	5	2.3	55	2	D59106	hypothetical prote
857	6	2.8	1948	2	S00485	gene 11-1 protein	930	5	2.3	55	2	AD1852	hypothetical prote
858	6	2.8	1972	2	A41604	myosin heavy chain	931	5	2.3	56	2	I40011	seppf protein - Bac
859	6	2.8	1972	2	JCS420	smooth muscle myos	932	5	2.3	56	2	T07222	hypothetical prote
860	6	2.8	2094	2	S33124	tpx protein - huma	933	5	2.3	58	2	H83721	hypothetical prote
861	6	2.8	2098	2	T18397	protein CTRP - mal	934	5	2.3	59	2	S35169	cytochrome P450 (c
862	6	2.8	2100	2	T18128	TL123.15 protein -	935	5	2.3	59	2	B82466	hypothetical prote
863	6	2.8	2109	1	I50421	aggreacan precursor	936	5	2.3	59	2	D82459	hypothetical prote
864	6	2.8	2110	2	H96803	unknown protein T5	937	5	2.3	59	2	I40012	hypothetical prote
865	6	2.8	2130	2	AB0821	probable exported	938	5	2.3	59	2	E70581	hypothetical prote
866	6	2.8	2204	2	A70524	probable PPE prote	939	5	2.3	59	2	B97654	hypothetical prote
867	6	2.8	2262	2	T30890	calcium channel al	940	5	2.3	59	2	D98151	hypothetical prote
868	6	2.8	2265	1	FNBO	fibronectin - bovi	941	5	2.3	60	2	A82462	hypothetical prote
869	6	2.8	2282	2	T42717	DNA-binding protei	942	5	2.3	60	2	D72735	hypothetical prote
870	6	2.8	2301	1	GNNYTN	genome polypotein	943	5	2.3	60	2	AC1948	hypothetical prote
871	6	2.8	2303	1	GNNYTN	genome polypotein	944	5	2.3	60	2	AE2387	hypothetical prote
872	6	2.8	2303	1	GNNYTP	genome polypotein	945	5	2.3	61	2	B83223	hypothetical prote
873	6	2.8	2303	1	S13554	genome polypotein	946	5	2.3	61	2	C84029	hypothetical prote
874	6	2.8	2383	2	D64962	probable membrane	947	5	2.3	61	2	A97764	hypothetical prote
875	6	2.8	2401	2	T28676	rioptry protein -	948	5	2.3	61	2	AG2122	hypothetical prote
876	6	2.8	2415	1	A33733	spectrin alpha cha	949	5	2.3	63	2	A83899	hypothetical prote
877	6	2.8	2663	1	S28261	centromere protein	950	5	2.3	63	2	AE3404	hypothetical prote
878	6	2.8	2688	2	I49477	alpha-A-crystallin	951	5	2.3	64	1	WMVZK6	IKK protein - vacci
879	6	2.8	2717	2	A34203	DNA-binding protei	952	5	2.3	64	2	T42109	replication-associ
880	6	2.8	2833	2	A43360	inositol 1,4,5-tri	953	5	2.3	64	2	S52615	hypothetical prote
881	6	2.8	2893	2	A64556	toxin-like outer m	954	5	2.3	65	2	E91196	hypothetical prote
882	6	2.8	3011	1	GNMVCH	genome polypotein	955	5	2.3	65	2	E84494	hypothetical prote
883	6	2.8	3016	2	S77300	hypothetical prote	956	5	2.3	65	2	T48279	hypothetical prote
884	6	2.8	3036	2	T18995	hypothetical prote	957	5	2.3	65	2	C82542	hypothetical prote
885	6	2.8	3110	2	AC0116	probable virulence	958	5	2.3	65	2	B85601	hypothetical prote
886	6	2.8	3573	2	S23070	erythronolide synt	959	5	2.3	65	2	D86603	unknown protein en
887	6	2.8	3890	2	C89921	hypothetical prote	960	5	2.3	65	2	F97134	hypothetical prote
888	6	2.8	3951	1	VFIH81	F1 protein - avian	961	5	2.3	65	2	AB1911	hypothetical prote
889	6	2.8	4096	2	A57039	DNA-activated prot	962	5	2.3	66	2	T17909	hypothetical prote
890	6	2.8	6642	2	T29757	protein UNC-89 - C	963	5	2.3	66	2	H81804	hypothetical prote
891	6	2.8	6805	2	S20901	titin - rabbit (fr	964	5	2.3	66	2	B81061	hypothetical prote
892	6	2.8	26926	1	I38344	titin, cardiac mus	965	5	2.3	67	2	A70375	ribosomal protein
893	5	2.3	16	2	A48301	glutamate-1-semial	966	5	2.3	67	2	JH0117	protein-tyrosine k
894	5	2.3	18	2	I40062	shikimate 5-dehydr	967	5	2.3	67	2	A99352	partial ORF from I
895	5	2.3	18	2	A30541	F7-1 fibrillar prot	968	5	2.3	67	2	D87032	conserved hypotet
896	5	2.3	20	2	A40451	domancy-related p	969	5	2.3	67	2	T10153	hypothetical prote
897	5	2.3	21	2	S09517	prolamin - sorghum	970	5	2.3	67	2	H98019	hypothetical prote
898	5	2.3	25	2	PH1910	T-cell receptor al	971	5	2.3	67	2	A10738	hypothetical prote
899	5	2.3	25	2	H41606	homeotic proteins	972	5	2.3	67	2	AF1982	conserved hypotet
900	5	2.3	26	2	I45087	cysteine proteinas	973	5	2.3	69	2	I46677	Nf1t protein [limp
901	5	2.3	30	2	C81791	hypothetical prote	974	5	2.3	69	2	P90926	collagen alpha 2(I
902	5	2.3	31	2	H72808	gp82.2 protein - M	975	5	2.3	69	2	B85775	hypothetical prote
903	5	2.3	33	2	PS0433	dyctrophin - rabbi	976	5	2.3	69	2	C64925	hypothetical prote
904	5	2.3	34	2	AB3789	hypothetical prote	977	5	2.3	69	2	AC1812	hypothetical prote
905	5	2.3	37	2	C84210	hypothetical prote	978	5	2.3	70	2	S07510	gene 4.3 protein -

```

579 5 2.3 70 2 E70985 hypothetical prote
580 5 2.3 70 2 I83374 protein-tyrosine k
581 5 2.3 70 2 A80700 conserved hypothet
582 5 2.3 70 2 A61591 hypothetical prote
583 5 2.3 71 2 A53533 glucose-6-phosphat
584 5 2.3 71 2 A60014 50S ribosomal prot
585 5 2.3 71 2 S01739 hypothetical prote
586 5 2.3 71 2 G82565 hypothetical prote
587 5 2.3 71 2 A82179 hypothetical prote
588 5 2.3 72 1 OAECD8 ydcF protein - Esc
589 5 2.3 72 1 B85608 hypothetical prote
590 5 2.3 73 2 D84810 probable proteinase
591 5 2.3 73 2 H71149 hypothetical prote
592 5 2.3 73 2 A82468 secretory protein
593 5 2.3 74 2 B82939 hypothetical prote
594 5 2.3 74 2 H75199 hypothetical prote
595 5 2.3 75 2 H72236 conserved hypothet
596 5 2.3 75 2 C90319 ort in partial tra
597 5 2.3 75 2 A82900 hypothetical prote
598 5 2.3 75 2 A80392 probable exported
599 5 2.3 76 2 S60827 M protein precursor
1000 5 2.3 76 2 D82844 carbon storage reg

```

## ALIGNMENTS

## RESULT 1

outer surface protein C precursor - Borrelia garinii

C/Species: Borrelia garinii  
C/Date: 04-Sep-1997 #sequence\_revision 04-Sep-1997 #text\_change 26-May-2000

C/Accession: I40272

R/Pukunaga, M.; Hamase, A.  
U. Clin. Microbiol. 33, 2415-2420, 1995

A/Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu la

A/Reference number: I40269; MUID:96025162; PMID:7494039

A/Accession: I40272

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA  
A/Residues: 1-210 <RES>

A/Cross-references: GB:D49378; NID:G1041105; PIDN:BA08376.1; PID:G1041106

C/Superfamily: Lyme disease spirochete surface protein C

```

Query Match          3.7%; Score 8; DB 2; Length 210;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 38 LAKEATLK 45

Db 154 LAKEATLK 161

RESULT 2  
G84608

En/Spm-like transposon protein [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C/Accession: G84608

R/Li, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J

A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A/Reference number: A84420; MUID:20083487; PMID:10617197

A/Accession: G84608

A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-140 <STO>

A/Cross-references: GB:AE002093; NID:G4587588; PIDN:AA025816.1; GSPDB:GN00139

C/Genetics: A:Gene: At2g22080  
A:Map position: 2

```

Query Match          3.2%; Score 7; DB 2; Length 140;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 155 ADDABGK 161

Db 46 ADDABGK 52

## RESULT 3

outer surface protein C - Lyme disease spirochete (strain TXGW) (fragment)

C/Species: Borrelia burgdorferi (Lyme disease spirochete)

C/Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 26-May-2000

C/Accession: I40129; S54199

R/Thiesen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.

U. Bacteriol. 177, 3036-3044, 1995

A/Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.

A/Reference number: I40104; MUID:95286481; PMID:7768799

A/Accession: I40129

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA  
A/Residues: 1-177 <RES>

A/Cross-references: EMBL:X84783; NID:G793825; PIDN:CAA59254.1; PID:G793826

C/Genetics: A:Gene: ospC  
C/Superfamily: Lyme disease spirochete surface protein C

```

Query Match          3.2%; Score 7; DB 2; Length 177;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 39 AKEATLK 45

Db 129 AKEATLK 135

## RESULT 4

outer surface protein C - Lyme disease spirochete

C/Species: Borrelia burgdorferi (Lyme disease spirochete)

C/Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 26-May-2000

C/Accession: S70262

R/Livley, I.; Gibbs, C.P.; Schuster, R.; Dornier, F.

Mol. Microbiol. 18, 257-269, 1995

A/Title: Evidence for lateral transfer and recombination in OspC variation in Lyme dise

A/Reference number: S70255; MUID:96296448; PMID:8709845

A/Accession: S70262

A/Status: nucleic acid sequence not shown

A/Molecule type: DNA  
A/Residues: 1-191 <LIV>

A/Cross-references: EMBL:L42879; NID:G858706; PIDN:AA836988.1; PID:G1695206

A/Experimental source: strain M57

C/Genetics: A:Gene: ospC  
C/Superfamily: Lyme disease spirochete surface protein C

```

Query Match          3.2%; Score 7; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 39 AKEATLK 45

Db 136 AKEATLK 142

## RESULT 5

outer surface protein C - Lyme disease spirochete

C/Species: Borrelia burgdorferi (Lyme disease spirochete)

C/Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 26-May-2000

C/Accession: S70272



R, Lively, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.  
Mol. Microbiol. 18, 257-269, 1995  
A>Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disease  
A:Reference number: S70255; MUID:96296448; PMID:8709845  
A:Accession: S70272  
A>Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-192 <LIV>  
A:Cross-references: EMBL:J42869; NID:G858736; PIDN:AA37012.1; PID:G1695227  
A:Experimental source: strain W  
C:Genetics:  
A:Gene: ospC  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 3.2%; Score 7; DB 2; Length 192;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 AKEALIK 45  
|||||  
Db 137 AKEALIK 143

RESULT 6  
S54196  
outer surface protein C - Borrelia garinii (strain SL10) (fragment)  
C:Species: Borrelia garinii  
C>Date: 08-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 08-Dec-2000  
A:Accession: I40127; S54196  
R:Reisen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.  
J. Bacteriol. 177, 3036-3044, 1995  
A>Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.  
A:Reference number: I40104; MUID:95286481; PMID:7768799  
A:Accession: I40127  
A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-196 <THB>  
A:Cross-references: EMBL:X84780; NID:G793821; PIDN:CAA59251.1; PID:G793822  
A>Note: submitted to the EMBL Data Library, February 1995  
C:Genetics:  
A:Gene: ospC  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 3.2%; Score 7; DB 2; Length 196;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 AKEALIK 45  
|||||  
Db 148 AKEALIK 154

RESULT 7  
C82619  
Heti protein XFI934 [imported] - Xylella fastidiosa (strain 9a5c)  
C:Species: Xylella fastidiosa  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
A:Accession: C82619  
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen  
Nature 406, 151-157, 2000  
A>Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: A82515; MUID:20365717; PMID:10910347  
A>Note: for a complete list of authors see reference number A59328 below  
A:Accession: C82619  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-197 <STM>  
A:Cross-references: GB:AE004013; GB:AE003849; NID:G9107030; PIDN:AF64736.1; GSPDB:GN001  
R:Simpan, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Agencio, M.; Alvares, R.; A  
Briener, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carre, H  
de-Neco, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.  
submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Praga, J.S.; Franca, S.C.; Franco, M.C.; Froh  
J.D.; Jungueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig  
chad, M.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins,  
A:Authors: Martins, E.M.F.; Matsushima, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.  
Rodrigues, V.; Rosa, A.U. de M.; de Rosa Ur., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasa  
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silve  
M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.;  
A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: XFI934

Query Match 3.2%; Score 7; DB 2; Length 197;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 AKEALIK 45  
|||||  
Db 130 AKEALIK 136

RESULT 8  
AC3318  
sodium-dependent phosphate transport protein [imported] - Brucella melitensis (strain 1  
C:Species: Brucella melitensis  
C>Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 01-Feb-2002  
A:Accession: AC3318  
R:Delvecchio, V.G.; Kapral, V.; Redkar, R.U.; Patra, G.; Mujer, C.; Ios, T.; Ivanova,  
.; Mazur, M.; Goldsman, B.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A>Title: The genome sequence of the facultative intracellular pathogen Brucella melitem  
A:Reference number: AD3552; PMID:11756688  
A:Accession: AC3318  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-202 <KOR>  
A:Cross-references: GB:AB008917; PIDN:ALU51710.1; PID:G17982445; GSPDB:GN00190  
A:Experimental source: strain 16M  
C:Genetics:  
A:Gene: BMEI0529  
A:Map position: 1

Query Match 3.2%; Score 7; DB 2; Length 202;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 AAVILIT 123  
|||||  
Db 186 AAVILIT 192

RESULT 9  
I40111  
outer surface protein C - Borrelia garinii (strain DK29) (fragment)  
C:Species: Borrelia garinii  
C>Date: 04-Sep-1997 #sequence\_revision 07-Nov-1997 #text\_change 08-Dec-2000  
A:Accession: I40111  
R:Reisen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.  
J. Bacteriol. 177, 3036-3044, 1995  
A>Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.  
A:Reference number: I40104; MUID:95286481; PMID:7768799  
A:Accession: I40111  
A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-203 <THE>  
A:Cross-references: EMBL:X84770; NID:G793795; PIDN:CAA59241.1; PID:G806330  
A>Note: submitted to the EMBL Data Library, February 1995  
C:Genetics:  
A:Gene: ospC  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 3.2%; Score 7; DB 2; Length 203;



Best Local Similarity 100.0%; Pred. No. 28;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 39 AKEAIIK 45  
|||||  
Db 155 AKEAIIK 161

## RESULT 10

140112  
outer surface protein C - Borrelia garinii (strain DK32) (fragment)  
C:Species: Borrelia garinii  
C:Date: 04-Sep-1997 #sequence\_revision 07-Nov-1997 #text\_change 08-Dec-2000  
C:Accession: 140112  
R:Thiesen, M.; Borre, M.; Mathiesen, M.J.; Mikkelsen, B.; Lebech, A.M.; Hansen, K.  
J. Bacteriol. 177, 3036-3044, 1995  
A:Title: Evolution of the Borrelia burgdorferi outer surface protein OspC.  
A:Reference number: 140104; MUID:95286481; PMID:7768799  
A:Accession: 140112  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-204 <THE>  
A:Cross-references: EMBL:X84772; NID:9793796; PIDN:CA59243.1; PID:9806331  
A:Note: Submitted to the EMBL Data Library, February 1995  
C:Genetics:  
A:Gene: ospC  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 3.2%; Score 7; DB 2; Length 204;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 AKEAIIK 45  
|||||  
Db 156 AKEAIIK 162

## RESULT 11

140105  
outer surface protein C - Borrelia garinii (strain DK27) (fragment)  
C:Species: Borrelia garinii  
C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 08-Dec-2000  
C:Accession: 140105; S34174  
R:Thiesen, M.; Frederiksen, B.; Lebech, A.M.; Vuust, J.; Hansen, K.  
J. Clin. Microbiol. 31, 2570-2576, 1993  
A:Title: Polymorphism in ospC gene of Borrelia burgdorferi and immunoreactivity of OspC  
A:Reference number: 140105; MUID:94075528; PMID:8253951  
A:Accession: 140105  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-205 <THE>  
A:Cross-references: EMBL:X73623; NID:9313271; PIDN:CA52002.1; PID:9313272  
A:Note: Submitted to the EMBL Data Library, February 1995  
C:Genetics:  
A:Gene: ospC  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 3.2%; Score 7; DB 2; Length 205;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 AKEAIIK 45  
|||||  
Db 157 AKEAIIK 163

## RESULT 12

S69926  
outer surface protein C precursor - Borrelia garinii (strain N34)  
C:Species: Borrelia garinii  
A:Variety: strain N34  
C:Date: 15-Feb-1997 #sequence\_revision 27-Feb-1997 #text\_change 26-May-2000  
C:Accession: S69926; S72680

R.Jauris-Heipke, S.; Liegl, G.; Preac-Mursic, V.; Roesler, D.; Schwab, E.; Soutschek, J. Clin. Microbiol. 33, 1860-1866, 1995  
A:Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borrelia  
A:Reference number: 140047; MUID:95395018; PMID:7665660  
A:Accession: S69926  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-209 <JAU>  
A:Cross-references: EMBL:X83556  
A:Experimental source: strain N34

R.Roesler, D.  
submitted to the EMBL Data Library, January 1995  
A:Reference number: S72679  
A:Accession: S72680  
A:Molecule type: DNA  
A:Residues: 1-173, 'D', 175-209 <ROE>  
A:Cross-references: EMBL:X83556; NID:9872023; PIDN:CA58546.1; PID:9872024  
C:Genetics:  
A:Gene: ospC  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 3.2%; Score 7; DB 2; Length 209;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 AKEAIIK 45  
|||||  
Db 154 AKEAIIK 160

## RESULT 13

S69920  
outer surface protein C precursor - Borrelia garinii (strain Phel)  
C:Species: Borrelia garinii  
A:Variety: strain Phel  
C:Date: 06-Dec-1996 #sequence\_revision 14-Feb-1997 #text\_change 26-May-2000  
C:Accession: S69920  
R.Jauris-Heipke, S.; Liegl, G.; Preac-Mursic, V.; Roesler, D.; Schwab, E.; Soutschek, J. Clin. Microbiol. 33, 1860-1866, 1995  
A:Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borrelia  
A:Reference number: 140047; MUID:95395018; PMID:7665660  
A:Accession: S69920  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-210 <JAU>  
A:Cross-references: EMBL:X83553; NID:9872025; PIDN:CA58543.1; PID:9872026  
A:Experimental source: strain Phel  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1995  
C:Genetics:  
A:Gene: ospC  
C:Superfamily: Lyme disease spirochete surface protein C  
F.1-18/Domain: signal sequence #status predicted <SIG>  
F.19-210/Product: outer surface protein C #status predicted <MAT>

Query Match 3.2%; Score 7; DB 2; Length 210;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 AKEAIIK 45  
|||||  
Db 155 AKEAIIK 161

## RESULT 14

S69925  
outer surface protein C precursor - Borrelia garinii (strain WABSou)  
C:Species: Borrelia garinii  
A:Variety: strain WABSou  
C:Date: 15-Feb-1997 #sequence\_revision 27-Feb-1997 #text\_change 26-May-2000  
C:Accession: S69925; S72678  
R.Jauris-Heipke, S.; Liegl, G.; Preac-Mursic, V.; Roesler, D.; Schwab, E.; Soutschek, J. Clin. Microbiol. 33, 1860-1866, 1995  
A:Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borrelia

A:Reference number: 140047; MUID:95395018; PMID:7665660  
A:Accession: S69925  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-210 <JAU>  
A:Cross-references: EMBL:X81526  
A:Experimental source: strain WABSON  
R:Roessler, D.  
Submitted to the EMBL Data Library, September 1994  
A:Reference number: S72674  
A:Accession: S72678  
A:Molecule type: DNA  
A:Residues: 1-86, 'D', 88-129, 'D', 131-210 <ROE>  
A:Cross-references: EMBL:X81526; NID:9804964; PIDN:CAA57246.1; PID:9804965  
C:Genetics:  
A:Gene: ospC  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 3.2%; Score 7; DB 2; Length 210;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 39 AKEAILK 45  
|||  
Db 155 AKEAILK 161

RESULT 15  
140274  
outer surface protein C precursor - Borrelia sp.  
C:Species: Borrelia sp.  
C>Date: 04-Sep-1997 #sequence\_revision 04-Sep-1997 #text\_change 26-May-2000  
C:Accession: 140274  
R:Fukunaga, M.; Hamase, A.  
J. Clin. Microbiol. 33, 2415-2420, 1995  
A:Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu la  
A:Reference number: 140269; MUID:96025162; PMID:7494039  
A:Accession: 140274  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-210 <RES>  
A:Cross-references: GB:D49381; NID:g1041109; PIDN:BA08379.1; PID:g1041110  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 3.2%; Score 7; DB 2; Length 210;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 39 AKEAILK 45  
|||  
Db 155 AKEAILK 161

RESULT 16  
140280  
outer surface protein C precursor - Borrelia sp.  
C:Species: Borrelia sp.  
C>Date: 04-Sep-1997 #sequence\_revision 04-Sep-1997 #text\_change 26-May-2000  
C:Accession: 140280  
R:Fukunaga, M.; Hamase, A.  
J. Clin. Microbiol. 33, 2415-2420, 1995  
A:Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu la  
A:Reference number: 140269; MUID:96025162; PMID:7494039  
A:Accession: 140280  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-210 <RES>  
A:Cross-references: GB:D49504; NID:g707099; PIDN:BA08464.1; PID:g769691  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 3.2%; Score 7; DB 2; Length 210;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 39 AKEAILK 45  
|||  
Db 155 AKEAILK 161

RESULT 17  
G70218  
outer surface protein C - Lyme disease spirochete  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C>Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 21-Jul-2000  
C:Accession: G70218; 140269; S37726; S70281  
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Whitson, D.; Peterson, J.; Kexlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vglt; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.  
Nature 390, 580-586, 1997  
A:Authors: Smith, H.O.; Venter, J.C.  
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.  
A:Reference number: A70100; MUID:98065943; PMID:9403685  
A:Accession: G70218  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-210 <KLB>  
A:Cross-references: GB:AE00792; NID:g3253098; PIDN:AAC66329.1; PID:g2689901; TIGR:BBB1  
A:Experimental source: strain B31  
R:Fukunaga, M.; Hamase, A.  
J. Clin. Microbiol. 33, 2415-2420, 1995  
A:Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu l  
A:Reference number: 140269; MUID:96025162; PMID:7494039  
A:Accession: 140269  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-210 <RES>  
A:Cross-references: GB:D49497; NID:g707092; PIDN:BA08457.1; PID:g769684  
R:Jauris-Haipre, S.; Fuchs, R.; Metz, M.; Preac-Mursic, V.; Schwab, E.; Soutchek, E.; Med. Microbiol. Immunol. 182, 37-50, 1993  
A:Title: Genetic heterogeneity of the genes coding for the outer surface protein C (ospC  
A:Reference number: S37726; MUID:93268136; PMID:8098841  
A:Accession: S37726  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-210 <JAU>  
A:Cross-references: EMBL:X69596; NID:g311391; PIDN:CAA49306.1; PID:g311392  
R:Livey, I.; Gibbs, C.P.; Schuster, R.; Dörner, F.  
Mol. Microbiol. 18, 257-269, 1995  
A:Title: Evidence for lateral transfer and recombination in ospC variation in Lyme dise  
A:Reference number: S70255; MUID:96296448; PMID:8709845  
A:Accession: S70281  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 19-210 <LIV>  
A:Cross-references: EMBL:L42887; NID:g858715; PIDN:AAB36995.1; PID:g1695212  
C:Experimental source: strain Ip2  
C:Genetics:  
A:Gene: ospC  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 3.2%; Score 7; DB 2; Length 210;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 39 AKEAILK 45  
|||  
Db 155 AKEAILK 161

RESULT 18  
S69927  
outer surface protein C precursor - Lyme disease spirochete (strain PKA)  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
A:Variety: strain PKA  
C>Date: 15-Feb-1997 #sequence\_revision 27-Feb-1997 #text\_change 26-May-2000  
C:Accession: S69927; S72669

R:Jauris-Heipke, S.; Liegl, G.; Preac-Mursic, V.; Roeseler, D.; Schwab, E.; Scutschek, H.  
J. Clin. Microbiol. 33, 1860-1866, 1995  
A:Title: Molecular analysis of genes encoding outer surface protein C (OspC) of *Borrelia*  
A:Reference number: 140047; PMID:95395018; PMID:7665660  
A:Accession: S69927  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-210 <JAU>  
A:Cross-references: EMBL:X69589  
A:Experimental source: strain PKA  
R:Jauris, S.  
submitted to the EMBL Data Library, February 1994  
A:Reference number: S72669  
A:Accession: S72669  
A:Molecule type: DNA  
A:Residues: 1-124, 'D', 126-139, 'E', 141-210 <JAU>  
A:Cross-references: EMBL:X69589  
C:Genetics:  
A:Gene: ospC  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 3.2%; Score 7; DB 2; Length 210;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 AKEALIK 45  
DB 155 AKEALIK 161

RESULT 19  
140275  
outer surface protein C precursor - *Borrelia garinii*  
C:Species: *Borrelia garinii*  
C:Date: 04-Sep-1997 #sequence\_revision 04-Sep-1997 #text\_change 26-May-2000  
C:Accession: 140275  
R:Fukunaga, M.; Hamase, A.  
J. Clin. Microbiol. 33, 2415-2420, 1995  
A:Title: Outer surface protein C gene sequence analysis of *Borrelia burgdorferi sensu la*  
A:Reference number: 140263; PMID:96025162; PMID:7494039  
A:Accession: 140275  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-212 <RES>  
A:Cross-references: GB:D49499; NID:9707094; PIDN:BAA08459.1; PID:9769686  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 3.2%; Score 7; DB 2; Length 212;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 AKEALIK 45  
DB 157 AKEALIK 163

RESULT 20  
B83003  
glutamine amidotransferase PA5142 [imported] - *Pseudomonas aeruginosa* (strain PA01)  
C:Species: *Pseudomonas aeruginosa*  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: B83003  
R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bz  
aman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kass, A.; Larbig, K.; Lam,  
; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho  
A:Reference number: A82950; PMID:20437337; PMID:10984043  
A:Accession: B83003  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-213 <STO>  
A:Cross-references: GB:AE004927; GB:AE004091; NID:9951437; PIDN:AAG08527.1; GSPDB:GN001

A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: hlyH1; PA5142  
C:Superfamily: amidotransferase hlyH; trpG homology

Query Match 3.2%; Score 7; DB 2; Length 213;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 179 GLOLLON 185  
DB 199 GLOLLON 205

RESULT 21  
S69916  
outer surface protein C precursor - *Borrelia afzelii* (strain Plud)  
C:Species: *Borrelia afzelii*  
A:Variety: strain Plud  
C:Date: 06-Dec-1996 #sequence\_revision 14-Feb-1997 #text\_change 26-May-2000  
C:Accession: S69916  
R:Jauris-Heipke, S.; Liegl, G.; Preac-Mursic, V.; Roeseler, D.; Schwab, E.; Scutschek,  
J. Clin. Microbiol. 33, 1860-1866, 1995  
A:Title: Molecular analysis of genes encoding outer surface protein C (OspC) of *Borrelia*  
A:Reference number: 140047; PMID:95395018; PMID:7665660  
A:Accession: S69916  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-214 <JAU>  
A:Cross-references: EMBL:X83552; NID:9872017; PIDN:CA58542.1; PID:9872018  
A:Experimental source: strain Plud  
C:Superfamily: Lyme disease spirochete surface protein C

Query Match 3.2%; Score 7; DB 2; Length 214;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 AKEALIK 45  
DB 159 AKEALIK 165

RESULT 22  
D95951  
probable two-component response regulator protein Smb21210 [imported] - *Sinorhizobium m*  
C:Species: *Sinorhizobium meliloti*  
C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001  
C:Accession: D95951  
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Herma  
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
A:Title: The complete sequence of the 1.603-kb pSymb megaplasmid from the N2-fixing end  
A:Reference number: A95842; PMID:21396508; PMID:11481431  
A:Accession: D95951  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-223 <KUR>  
A:Cross-references: GB:AL591985; PIDN:CA649276.1; PID:G15140762; GSPDB:GN00167  
A:Experimental source: strain 1021, megaplasmid pSymb  
R:Galibert, F.; Finan, T.M.; Long, S.R.; Fuhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler  
L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Leilaure  
hebaull, P.; Vanderbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K  
A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.  
A:Reference number: A96039; PMID:21368234; PMID:11474104  
A:Contents: annotation  
C:Genetics:  
A:Gene: Smb21210  
A:Genome: plasmid  
C:Superfamily: ompR protein; response regulator homology

Query Match 3.2%; Score 7; DB 2; Length 223;

Best Local Similarity 100.0%; Pred. No. 30;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 178 DGLQLIQ 184  
Db 58 DGLQLIQ 64

RESULT 23

T35767  
hypothetical protein SC7H2.36c - Streptomyces coelicolor

C/Species: Streptomyces coelicolor  
C/Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 08-Sep-2000  
C/Accession: T35767  
R/Sundera, D.C.; Harrie, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.;  
submitted to the EMBL Data Library, August 1999  
A/Reference number: Z21588  
A/Accession: T35767  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-252 <SAU>  
A/Cross-references: EMBL:AL109732; PIDN:CAB52078.1; GSPDB:GN00070; SCODEB:SC7H2.36c  
A/Experimental source: strain A3(2)  
A/Genetic: SCODEB:SC7H2.36c  
C/Superfamily: Streptomyces coelicolor hypothetical protein SC7H2.36c

Query Match 3.2%; Score 7; DB 2; Length 252;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 18 VDYSGSM 24  
Db 54 VDYSGSM 60

RESULT 24

C75077

from (111) abc transporter, ATP-binding protein (hemv-1) PAB0678 - Pyrococcus abyssi (at

C/Species: Pyrococcus abyssi  
C/Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 21-Jan-2000  
C/Accession: C75077  
R/anonymous, Genoscope  
submitted to the EMBL Data Library, July 1999  
A/Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru

A/Reference number: A75001  
A/Accession: C75077  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-257 <KAM>  
A/Cross-references: GB:AJ248286; GB:AL096836; NID:G5458366; PIDN:CAB49920.1; PID:el51581  
A/Experimental source: strain Orsay  
A/Genetic:

A/Gene: hemv-1; PAB0678  
C/Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology  
F/21-215/Domain: ATP-binding cassette homology <ABC>

Query Match 3.2%; Score 7; DB 2; Length 257;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 210 VLRGVNF 216  
Db 22 VLRGVNF 28

RESULT 25

G83178

conserved hypothetical protein PA3747 [imported] - Pseudomonas aeruginosa (strain PA01)

C/Species: Pseudomonas aeruginosa  
C/Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C/Accession: G83178  
R/Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path  
A/Reference number: A82950; MUID:20437337; PMID:10984043  
A/Accession: G83178  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-266 <STO>  
A/Cross-references: GB:AE004793; GB:AE004091; NID:G94949904; PIDN:AA07134.1; GSPDB:GN00

A/Experimental source: strain PA01  
A/Genetic: PA3747  
Query Match 3.2%; Score 7; DB 2; Length 266;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 117 AAVILLT 123  
Db 74 AAVILLT 80

RESULT 26

E83858

hypothetical protein BH1669 [imported] - Bacillus halodurans (strain C-125)

C/Species: Bacillus halodurans  
C/Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
C/Accession: E83858  
R/Takam, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hir  
Nucleic Acids Res. 28, 4317-4331, 2000  
A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
A/Reference number: A83650; MUID:20512582; PMID:11058132  
A/Accession: E83858  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-267 <STO>  
A/Cross-references: GB:AP001512; GB:BA000004; NID:G10174030; PIDN:BA05388.1; GSPDB:GN00

A/Experimental source: strain C-125  
A/Genetic: BH1669  
Query Match 3.2%; Score 7; DB 2; Length 267;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 57 GGLYTFA 63  
Db 66 GGLYTFA 72

RESULT 27

A84309

cobalamin biosynthesis [imported] - Halobacterium sp. NRC-1

C/Species: Halobacterium sp. NRC-1  
C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001  
C/Accession: A84309  
R/Ng, W.V.; Kennedy, S.P.; Mahairs, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky,  
.; Leitbauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A/Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; L  
A/Title: Genome sequence of Halobacterium species NRC-1.  
A/Reference number: A84160; MUID:20504483; PMID:11016550  
A/Accession: A84309  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-276 <STO>  
A/Cross-references: GB:AE004437; NID:G10581041; PIDN:AA019837.1; GSPDB:GN00138

A/Genetic: COB  
C/Superfamily: probable precorrin-3 methylase

Query Match 3.2%; Score 7; DB 2; Length 276;  
 Best Local Similarity 100.0%; Pred. No. 37;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 36 IELAXA 42  
 |||||  
 Db 78 IELAXA 84

## RESULT 28

probable ABC transporter, ATP binding protein Sma1120 [imported] - Sinorhizobium meliloti  
 C:Species: Sinorhizobium meliloti  
 C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 17-May-2002  
 C:Accession: F95337  
 R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows, K.; Kaiman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.; Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001  
 A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti  
 A:Reference number: A95262; MUID:21396509; PMID:11481432  
 A:Accession: F95337  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-304 <KUR>  
 A:Cross-references: GB:AE006469; PIDN:AAK65264.1; PID:G14523716; GSPDB:GN00165  
 A:Experimental source: strain 1021, megaplasmid pSymA  
 R:Galibert, F.; Flman, T.M.; Long, S.R.; Philier, A.; Abola, P.; Ampe, F.; Barloy-Hubler, P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T.  
 Science 293, 668-672, 2001  
 A:Author: Kahn, D.; Kahn, M.L.; Kaiman, S.; Keating, D.H.; Kias, E.; Komp, C.; Laure, H.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Welle, D.H.; Wong, K.; Yeh, K.  
 A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
 A:Reference number: A96039; MUID:21368234; PMID:11474104  
 A:Contents: annotation  
 C:Genetics:  
 A:Gene: Sma1120  
 A:Genome: plasmid  
 C:Superfamily: short-chain ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 3.2%; Score 7; DB 2; Length 304;  
 Best Local Similarity 100.0%; Pred. No. 40;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 41 EATIKIN 47  
 |||||  
 Db 250 EATIKIN 256

## RESULT 29

probable phosphatidate cytidyltransferase PA2536 [imported] - Pseudomonas aeruginosa  
 C:Species: Pseudomonas aeruginosa  
 C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 C:Accession: A83329  
 R:Stover, C.K.; Pham, X.Q.; Ewlin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bradman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kae, A.; Lattig, K.; Lim, N.; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen  
 A:Reference number: A82950; MUID:20437337; PMID:10984043  
 A:Accession: A83329  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-311 <STO>  
 A:Cross-references: GB:AE004691; GB:AE004091; NID:G9948587; PIDN:AA05924.1; GSPDB:GN001  
 A:Experimental source: strain PA01  
 C:Genetics:  
 A:Gene: PA2536

Query Match 3.2%; Score 7; DB 2; Length 311;  
 Best Local Similarity 100.0%; Pred. No. 41;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 116 QAAVILL 122  
 |||||  
 Db 63 QAAVILL 69

## RESULT 30

probable ABC-type transport protein abc - Escherichia coli (strain K-12)  
 C:Species: Escherichia coli  
 C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 01-Mar-2002  
 C:Accession: G64744; I4113  
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C.A.; Rose, D.J.; Mau, B.; Shao, Y.  
 Science 277, 1453-1462, 1997  
 A:Title: The complete genome sequence of Escherichia coli K-12.  
 A:Reference number: A64720; MUID:97426617; PMID:9278503  
 A:Accession: G64744  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-343 <BLAT>  
 A:Cross-references: GB:AE000129; GB:U00096; NID:G1786395; PIDN:AA073310.1; PID:G1786398  
 A:Experimental source: strain K-12, substrain MG1655  
 R:AlIkmet, R.; Gerard, B.; Court, D.; Dean, M.  
 Gene 136, 231-236, 1993  
 A:Title: Cloning and organization of the abc and mgl genes of Escherichia coli: relation  
 A:Reference number: I4113; MUID:94124004; PMID:7904973  
 A:Accession: I4113  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-164, 'V', 166-199, 'KMTL', <RES>  
 A:Cross-references: GB:L08626; NID:G145168; PIDN:AA036869.1; PID:G145169  
 C:Genetics:  
 A:Gene: abc  
 C:Superfamily: probable ABC-type transport protein abc; ATP-binding cassette homology  
 C:Keywords: ATP; nucleotide binding; P-loop  
 F:21-217/Domain: ATP-binding cassette homology <ABC>  
 F:38-45/Region: nucleotide-binding motif A (P-loop)

Query Match 3.2%; Score 7; DB 2; Length 343;  
 Best Local Similarity 100.0%; Pred. No. 45;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 173 GSVLVDG 179  
 |||||  
 Db 60 GSVLVDG 66

## RESULT 31

ATP-binding component of a transporter Eca0201 [imported] - Escherichia coli (strain O1A90654)  
 C:Species: Escherichia coli  
 C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001  
 C:Accession: A90654  
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Gessawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
 DNA Res. 8, 11-22, 2001  
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen  
 A:Reference number: A99629; MUID:21156231; PMID:11258796  
 A:Accession: A90654  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-343 <HAY>  
 A:Cross-references: GB:BA000007; PIDN:BA033624.1; PID:G13359657; GSPDB:GN00154  
 A:Experimental source: strain O157:H7, substrain RIMD 0509952  
 C:Genetics:  
 A:Gene: Eca0201  
 C:Superfamily: probable ABC-type transport protein abc; ATP-binding cassette homology

Query Match 3.2%; Score 7; DB 2; Length 343;  
 Best Local Similarity 100.0%; Pred. No. 45;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 GSVLVDG 179  
 |||||  
 Db 60 GSVLVDG 66

RESULT 32

ATP-binding component of a transporter [imported] - Escherichia coli (strain O157:H7, su  
 A:Accession: A85505  
 C:Species: Escherichia coli  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
 R:Accession: A85505  
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
 11ler, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoudis, K.; Apodaca,  
 Nature 409, 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
 A:Reference number: A85480; MUID:21074935; PMID:11206551  
 A:Accession: A85505  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-343 <STO>  
 A:Cross-references: GB:AE005174; NID:912512930; PIDN:AAG54501.1; GSPDB:GN00145; UWGP:202  
 A:Experimental source: strain O157:H7, substrain EDL933  
 C:Genetic8:  
 A:Gene: abc  
 C:Superfamily: probable ABC-type transport protein abc; ATP-binding cassette homology

Query Match 3.2% Score 7; DB 2; Length 343;  
 Best Local Similarity 100.0%; Pred. No. 45;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 GSVLVDG 179  
 |||||  
 Db 60 GSVLVDG 66

RESULT 33

conserved hypothetical protein CC1169 [imported] - Caulobacter crescentus  
 C:Species: Caulobacter crescentus  
 C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
 C:Accession: E87394  
 R:Neuman, W.C.; Felblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
 B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolor  
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.;  
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
 A:Title: Complete Genome Sequence of Caulobacter crescentus.  
 A:Reference number: A87249; MUID:21173698; PMID:11259647  
 A:Accession: E87394  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-355 <STO>  
 A:Cross-references: GB:AE005673; NID:91342491; PIDN:AAK2153.1; GSPDB:GN00148  
 C:Genetic8:  
 A:Gene: CC1169

Query Match 3.2% Score 7; DB 2; Length 355;  
 Best Local Similarity 100.0%; Pred. No. 46;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 182 LQNPVAV 188  
 |||||  
 Db 99 LQNPVAV 105

RESULT 34

phosphoglycerate kinase (EC 2.7.2.3) - Synechocystis sp. (strain PCC 6803)  
 C:Species: Synechocystis sp.  
 C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 20-Jun-2000  
 C:Accession: S76392  
 R:Kanehko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;  
 o, K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda

DNA Res. 3, 109-136, 1996  
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocysti  
 s.  
 A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S76392  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-374 <KAN>  
 A:Cross-references: EMBL:D90915; GB:AB001339; NID:91653604; PIDN:BA18521.1; PID:916536  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
 C:Genetic8:  
 A:Start codon: GTG  
 C:Superfamily: phosphoglycerate kinase  
 C:Keywords: phosphotransferase

Query Match 3.2% Score 7; DB 2; Length 374;  
 Best Local Similarity 100.0%; Pred. No. 49;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 RLTPVGD 101  
 |||||  
 Db 49 RLTPVGD 55

RESULT 35

T04608  
 ADP-ATP carrier protein F2009.60 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 21-May-1999  
 C:Accession: T04608  
 R:Bevan, M.; Rose, M.; Hempel, S.; Eutian, K.D.; Hohnsbeil, J.; Mewes, H.W.; Mayer, K.F.  
 submitted to the Protein Sequence Database, October 1998  
 A:Reference number: Z15380  
 A:Accession: T04608  
 A:Molecule type: DNA  
 A:Residues: 1-379 <BEV>  
 A:Cross-references: EMBL:AL021749  
 A:Experimental source: cultivar Columbia; BAC clone F2009  
 C:Genetic8:  
 A:Map position: 4  
 A:Introns: 157/3; 281/3  
 A:Note: F2009.60  
 C:Superfamily: ADP-ATP carrier protein; ADP-ATP carrier protein repeat homology  
 C:Keywords: duplication; transmembrane protein  
 F:76-171/Domain: ADP-ATP carrier protein repeat homology <ACP2>  
 F:181-275/Domain: ADP-ATP carrier protein repeat homology <ACP3>  
 F:281-369/Domain: ADP-ATP carrier protein repeat homology <ACP3>

Query Match 3.2% Score 7; DB 2; Length 379;  
 Best Local Similarity 100.0%; Pred. No. 49;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 175 VLVDGLQ 181  
 |||||  
 Db 275 VLVDGLQ 281

RESULT 36

hypothetical protein Z1487 [imported] - Escherichia coli (strain O157:H7, substrain EDL  
 C:Species: Escherichia coli  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
 C:Accession: D85643  
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
 11ler, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoudis, K.; Apodaca,  
 Nature 409, 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
 A:Reference number: A85480; MUID:21074935; PMID:11206551  
 A:Accession: D85643  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-407 <STO>  
 A:Cross-references: GB:AE005174; NID:912514343; PIDN:AAG55608.1; GSPDB:GN00145; UWGP:21

A:Experimental source: strain 0157:H7, substrain EDL933  
C:Genetics:  
A:Gene: Z1487

Query Match 3.2%; Score 7; DB 2; Length 407;  
Best Local Similarity 100.0%; Pred. No. 53;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 161 KAIIDQI 167  
|||||  
Db 153 KAIIDQI 159

## RESULT 37

A99783  
probable tail tip fiber protein [imported] - Escherichia coli (strain 0157:H7, substrain

C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001

C:Accession: A99783

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc  
A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: A99783

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-422 <HAY>

A:Cross-references: GB:BA000007; PIDN:BA934656.1; PID:G13360693; GSPDB:GN00154

A:Experimental source: strain 0157:H7, substrain RIMD 050952

C:Genetics:  
A:Gene: EC91233

Query Match 3.2%; Score 7; DB 2; Length 422;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 161 KAIIDQI 167  
|||||  
Db 166 KAIIDQI 174

## RESULT 38

E64617  
hypothetical protein HP0781 - Helicobacter pylori (strain 26695)

C:Species: Helicobacter pylori

C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999

C:Accession: E64617

R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.

Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenna

son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.

Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Frazer, C.

A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A:Reference number: A64520; MUID:97394467; PMID:9252185

A:Accession: E64617

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-429 <TOM>

A:Cross-references: GB:AE000590; GB:AE000511; NID:G2133907; PIDN:AAD07834.1; PID:G213391

Query Match 3.2%; Score 7; DB 2; Length 429;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 EPIKELA 39  
|||||  
Db 105 EPIKELA 111

## RESULT 39

F71896  
hypothetical protein jhp0718 - Helicobacter pylori (strain J99)

C:Species: Helicobacter pylori  
A:Variety: strain J99

C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 08-Oct-1999

C:Accession: F71896

R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.

7; Ives, C.; Gibson, R.; Werberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;

Nature 397, 176-180, 1999

A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric pat

A:Reference number: A71800; MUID:99120557; PMID:9923682

A:Accession: F71896

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-429 <ARN>

A:Cross-references: GB:AE001503; GB:AE001439; NID:G4155275; PIDN:AAD06301.1; PID:G41552

C:Genetics:  
A:Gene: jhp0718

Query Match 3.2%; Score 7; DB 2; Length 429;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 EPIKELA 39  
|||||  
Db 105 EPIKELA 111

## RESULT 40

AD2474  
hypothetical protein alr5348 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp.

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002

C:Accession: AD2474

R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Matanabe, A.; Iriuguch

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata,

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AD2474

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-447 <KUR>

A:Cross-references: GB:BA000019; PIDN:BA87047.1; PID:G17134487; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:  
A:Gene: alr5348

Query Match 3.2%; Score 7; DB 2; Length 447;  
Best Local Similarity 100.0%; Pred. No. 57;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 IELAKEA 42  
|||||  
Db 138 IELAKEA 144

Search completed: April 9, 2003, 14:46:03  
Job time : 73 secs





CC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxID=139;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DK27;  
RX MEDLINE=94075528; PubMed=8253951;  
RA Theisen M., Frederiksen B., Lebech A.M., Vuust J., Hansen K.;  
RT "Polymorphism in ospC gene of Borrelia burgdorferi and  
RT immunoreactivity of OspC protein: implications for taxonomy and for  
RT use of OspC protein as a diagnostic antigen.";  
RL J. Clin. Microbiol. 31:2570-2576(1993).  
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID  
ANCHOR.  
CC EMBL; X73623; CAA52002.1; -;  
DR InterPro; IPR001800; Lipoprotein\_6.  
DR Pfam; PF01441; Lipoprotein\_6; 1.  
DR ProDom; PD001149; Lipoprotein\_6; 1.  
FT SIGNAL; Lipoprotein; Outer membrane.  
FT SIGNAL 1 18  
FT CHAIN 19 205 POTENTIAL.  
FT NON\_TER 205 205 OUTER SURFACE PROTEIN C.  
SQ SEQUENCE 205 AA; 21426 MW; 3246712B96D2A510 CRC64;

Query Match 3.2%; Score 7; DB 2; Length 205;  
Best Local Similarity 100.0%; Pred. No. 59;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 39 AKEATLK 45  
Db 157 AKEATLK 163

RESULT 35  
O45179 PRELIMINARY; PRT; 209 AA.  
ID O45179  
AC O45179  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE Outer surface protein C precursor.  
GN OSPC.  
OS Borrelia garinii.  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxID=29519;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=N34;  
RX MEDLINE=95213332; PubMed=7699024;  
RA Wilske B., Jauris-Heipke S., Lobentanz R., Pradel I.,  
RA Preac-Mursec V., Rosales D., Soutschek E., Johnson R.C.;  
RT "Phenotypic analysis of outer surface protein C (OspC) of Borrelia  
RT burgdorferi sensu lato by monoclonal antibodies: relationship to  
RT genospecies and OspA serotype.";  
RL J. Clin. Microbiol. 33:103-109(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=N34;  
RX MEDLINE=95395018; PubMed=765660;  
RA Jauris-Heipke S., Diehl G., Preac-Mursec V., Roessler D., Schwab E.,  
RA Soutschek E., Wilske B.;  
RT "Molecular analysis of genes encoding outer surface protein C (OspC)  
RT of Borrelia burgdorferi sensu lato: relationship to ospA genotype and  
RT evidence of lateral gene exchange of ospC.";  
RL J. Clin. Microbiol. 33:1860-1866(1995).  
DR EMBL; X83556; CAA58546.1; -;  
DR InterPro; IPR001800; Lipoprotein\_6.  
DR Pfam; PF01441; Lipoprotein\_6; 1.  
DR ProDom; PD001149; Lipoprotein\_6; 1.  
FT SIGNAL 1 18  
FT CHAIN 19 209 POTENTIAL.  
FT NON\_TER 209 209 POTENTIAL.  
SQ SEQUENCE 209 AA; 22035 MW; 9DDB0164F9E4FC59 CRC64;

Query Match 3.2%; Score 7; DB 2; Length 209;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 39 AKEATLK 45  
Db 154 AKEATLK 160

RESULT 36  
O9KIK3 PRELIMINARY; PRT; 209 AA.  
ID O9KIK3  
AC O9KIK3  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Outer surface protein C.  
GN OSPC.  
OS Borrelia garinii.  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxID=29519;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K48;  
RA Dykhuizen D.E., Guttman D.S., Luft B.J.;  
RT "Antigenic variation and intragenic recombination in the ospC gene of  
RT Borrelia burgdorferi.";  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF231450; AAF76855.1; -;  
DR InterPro; IPR001800; Lipoprotein\_6.  
DR Pfam; PF01441; Lipoprotein\_6; 1.  
DR ProDom; PD001149; Lipoprotein\_6; 1.  
SQ SEQUENCE 209 AA; 21977 MW; C98F0164F9E4FAF5 CRC64;

Query Match 3.2%; Score 7; DB 2; Length 209;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 39 AKEATLK 45  
Db 154 AKEATLK 160

RESULT 37  
O9KIM6 PRELIMINARY; PRT; 210 AA.  
ID O9KIM6  
AC O9KIM6  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Outer surface protein C.  
GN OSPC.  
OS Borrelia garinii.  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxID=29519;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DK29;  
RA Dykhuizen D.E., Guttman D.S., Luft B.J.;  
RT "Antigenic variation and intragenic recombination in the ospC gene of  
RT Borrelia burgdorferi.";  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF230183; AAF75624.1; -;  
DR InterPro; IPR001800; Lipoprotein\_6.  
DR Pfam; PF01441; Lipoprotein\_6; 1.  
DR ProDom; PD001149; Lipoprotein\_6; 1.  
SQ SEQUENCE 210 AA; 22033 MW; A7760EB09AFAECB8 CRC64;

Query Match 3.2%; Score 7; DB 2; Length 210;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 39 AKEATLK 45

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=15M / ATCC 23456 / BIOTYPE 1;  
 RX MEDLINE=20020109; PubMed=11756688;  
 RA Delvecchio V.G., Kapatral V., Redkar R.J., Patra G., Muijer C., Los T.,  
 R Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,  
 RA Jablonksi L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,  
 RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Leeson J.-J.,  
 RA Haselhorn R., Kyriades N., Overbeek R.,  
 RT "The genome sequence of the facultative intracellular pathogen  
 RT Brucella melitensis."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).  
 DR EMBL; AB009495; AAL51710.1; -  
 DR InterPro; IPR003841; Na\_pl\_cotrans.  
 DR Pfam; PF02690; Na\_pl\_cotrans; 1.  
 KM Complete proteome.  
 SQ SEQUENCE 202 AA; 21749 MW; AE86F9F29AF6B662 CRC64;

Query Match 3.2%; Score 7; DB 16; Length 202;  
 Best Local Similarity 100.0%; Pred. No. 58;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 AAVILLT 123  
 |||||  
 DB 186 AAVILLT 192

## RESULT 31

Q44983 PRELIMINARY; PRT; 203 AA.

AC Q44983;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 DE Outer surface protein C (Fragment).  
 GN OSpC.  
 OS Borrelia garinii.  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=29519;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DK29;  
 RX MEDLINE=95286481; PubMed=7768799;  
 RA Theissen M., Borre M., Mathiesen M.J., Mikkelsen B., Lebech A.M.,  
 RA Hansen K.;  
 RT "Evolution of the Borrelia burgdorferi outer surface protein OSpC";  
 RL J. Bacteriol. 177:3036-3044(1995).  
 DR EMBL; X84770; CA59241.1; -  
 DR InterPro; IPR001800; Lipoprotein\_6.  
 DR Pfam; PF01441; Lipoprotein\_6; 1.  
 DR ProDom; PD001149; Lipoprotein\_6; 1.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 203 AA; 21226 MW; 25E1F4180847F7D4 CRC64;

Query Match 3.2%; Score 7; DB 2; Length 203;  
 Best Local Similarity 100.0%; Pred. No. 58;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 AKKAILK 45  
 |||||  
 DB 155 AKKAILK 161

## RESULT 32

Q44984 PRELIMINARY; PRT; 204 AA.

AC Q44984;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 DE Outer surface protein C (Fragment).  
 GN OSpC.

OS Borrelia garinii.  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=29519;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DK32;  
 RX MEDLINE=95286481; PubMed=7768799;  
 RA Theissen M., Borre M., Mathiesen M.J., Mikkelsen B., Lebech A.M.,  
 RA Hansen K.;  
 RT "Evolution of the Borrelia burgdorferi outer surface protein OSpC";  
 RL J. Bacteriol. 177:3036-3044(1995).  
 DR EMBL; X84772; CA59243.1; -  
 DR InterPro; IPR001800; Lipoprotein\_6.  
 DR Pfam; PF01441; Lipoprotein\_6; 1.  
 DR ProDom; PD001149; Lipoprotein\_6; 1.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 204 AA; 21369 MW; 1405DDCOA8B3D762 CRC64;

Query Match 3.2%; Score 7; DB 2; Length 204;  
 Best Local Similarity 100.0%; Pred. No. 59;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 AKKAILK 45  
 |||||  
 DB 156 AKKAILK 162

## RESULT 33

O50622 PRELIMINARY; PRT; 205 AA.

AC O50622;  
 DT 01-JUN-1998 (TREMblrel. 06, Created)  
 DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE Outer surface protein C (Fragment).  
 GN OSpC.  
 OS Borrelia afzelii.  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=29518;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PLUD;  
 RA Maenzawa T., Sawaki K., Yoshii T.;  
 RT "B. afzelii OSpC (Plud)." to the EMBL/GenBank/DBJ databases.  
 RL Submitted (DEC-1997)  
 DR EMBL; AB009898; BAA24128.1; -  
 DR InterPro; IPR001800; Lipoprotein\_6.  
 DR Pfam; PF01441; Lipoprotein\_6; 1.  
 DR ProDom; PD001149; Lipoprotein\_6; 1.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 205 AA; 21301 MW; 09741AB4048BB1C CRC64;

Query Match 3.2%; Score 7; DB 2; Length 205;  
 Best Local Similarity 100.0%; Pred. No. 59;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 AKKAILK 45  
 |||||  
 DB 153 AKKAILK 159

## RESULT 34

O08141 PRELIMINARY; PRT; 205 AA.

AC O08141;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 DE Outer surface protein C precursor (Fragment).  
 GN OSpC.  
 OS Borrelia burgdorferi (Lyme disease spirochete).

RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,  
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,  
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.,  
 RT "The genome sequence of the plant pathogen *Xylella fastidiosa*,"  
 RL Nature 406:151-159(2000).  
 DR EMBL: AB004013; AAF84736.1; -  
 DR InterPro: IPR002582; ACPS.  
 DR Pfam: PF01648; ACPS; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 197 AA; 22414 MW; 4936BAFD31902C17 CRC64;

Query Match 3.2%; Score 7; DB 16; Length 197;  
 Best Local Similarity 100.0%; Pred. No. 57;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 AKEATLK 45  
 |||||  
 DB 130 AKEATLK 136

## RESULT 27

Q9R0R9 PRELIMINARY; PRT; 200 AA.  
 AC Q9R0R9;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Outer surface protein C (Fragment).  
 GN OSPC.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=2E7;  
 RX MEDLINE=20002545; PubMed=10531219;  
 RA Hottelster E.K., Glass G.E., Childs J.E., Persing D.H.,  
 RT "Population dynamics of a naturally occurring heterogeneous mixture of  
 RT Borrelia burgdorferi clones,"  
 RL Infect. Immun. 67:5709-5716(1999).  
 DR EMBL: AF074464; AAD23911.1; -  
 DR InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF01441; Lipoprotein\_6; 1.  
 DR ProDom: PD001149; Lipoprotein\_6; 1.  
 KW Plasmid.  
 FT NON\_TER  
 SQ SEQUENCE 200 AA; 21307 MW; A22F158758BBB6B CRC64;

Query Match 3.2%; Score 7; DB 2; Length 200;  
 Best Local Similarity 100.0%; Pred. No. 57;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 AKEATLK 45  
 |||||  
 DB 155 AKEATLK 161

## RESULT 28

Q93Q98 PRELIMINARY; PRT; 202 AA.  
 ID Q93Q98;  
 AC Q93Q98;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Outer surface protein C (Fragment).  
 GN OSPC.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MI-2;

RA Lin T., Oliver J.H., Gao L.,  
 RT "Outer surface protein C gene sequence analysis of *Borrelia*  
 RT burgdorferi sensu lato isolated in the Southern United States,"  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF278582; AAK69455.1; -  
 DR InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF01441; Lipoprotein\_6; 1.  
 DR ProDom: PD001149; Lipoprotein\_6; 1.  
 FT NON\_TER  
 SQ SEQUENCE 202 AA; 21426 MW; D196E57979D698B6 CRC64;

Query Match 3.2%; Score 7; DB 2; Length 202;  
 Best Local Similarity 100.0%; Pred. No. 58;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 AKEATLK 45  
 |||||  
 DB 149 AKEATLK 155

## RESULT 29

Q932W2 PRELIMINARY; PRT; 202 AA.  
 ID Q932W2;  
 AC Q932W2;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Outer surface protein C (Fragment).  
 GN OSPC.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=B31, BBDM, AND BBDP;  
 RA Lin T., Oliver J.H., Gao L.,  
 RT "Outer surface protein C gene sequence analysis of *Borrelia*  
 RT burgdorferi sensu lato isolated in the Southern United States,"  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF278578; AAK69451.1; -  
 DR EMBL: AF278579; AAK69452.1; -  
 DR EMBL: AF278580; AAK69453.1; -  
 DR EMBL: AF278581; AAK69454.1; -  
 DR InterPro: IPR001800; Lipoprotein\_6.  
 DR Pfam: PF01441; Lipoprotein\_6; 1.  
 DR ProDom: PD001149; Lipoprotein\_6; 1.  
 FT NON\_TER  
 SQ SEQUENCE 202 AA; 21399 MW; CBBF57979D698B6 CRC64;

Query Match 3.2%; Score 7; DB 2; Length 202;  
 Best Local Similarity 100.0%; Pred. No. 58;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 AKEATLK 45  
 |||||  
 DB 149 AKEATLK 155

## RESULT 30

Q8Y1B5 PRELIMINARY; PRT; 202 AA.  
 ID Q8Y1B5;  
 AC Q8Y1B5;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Sodium-dependent phosphate transport protein.  
 GN BME10529.  
 OS Brucella melitensis.  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Brucellaceae; Brucella.  
 OX NCBI\_TaxID=29459;

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE OepC (Fragment).  
GN OepC.  
OS Borrelia garinii.  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxId=29519;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=FA803;  
RA Aas Gylfe B., Olsen B., Strasevicius D., Marti Ras N., Weihe P.,  
RA Noppa L., Ostberg Y., Baranton G., Bergstrom S.,  
RT " Lyme disease Borrelia isolated from puffins (Fratercula arctica) and  
RT the seabird ticks (Ixodes uriae) on the Faeroe Islands.",  
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF080263; AAD4501.1; -  
DR InterPro; IPR01800; Lipoprotein\_6.  
DR Pfam; PF01441; Lipoprotein\_6; 1.  
DR Prodom; PD01149; Lipoprotein\_6; 1.  
FT NON\_TER 1  
FT SEQUENCE 193 AA; 20350 MW; FB8AD969D89661CE CRC64;  
SQ  
Query Match 3.2%; Score 7; DB 2; Length 193;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 39 AKAILK 45  
DB 145 AKAILK 151

RESULT 24  
Q9S516 PRELIMINARY; PRT; 194 AA.  
AC Q9S516;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE OepC (Fragment).  
GN OepC.  
OS Borrelia garinii.  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxId=29519;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=FA802;  
RA Aas Gylfe B., Olsen B., Strasevicius D., Marti Ras N., Weihe P.,  
RA Noppa L., Ostberg Y., Baranton G., Bergstrom S.,  
RT " Lyme disease Borrelia isolated from puffins (Fratercula arctica) and  
RT the seabird ticks (Ixodes uriae) on the Faeroe Islands.",  
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF080264; AAD4502.1; -  
DR InterPro; IPR01800; Lipoprotein\_6.  
DR Pfam; PF01441; Lipoprotein\_6; 1.  
DR Prodom; PD01149; Lipoprotein\_6; 1.  
FT NON\_TER 1  
FT NON\_TER 194  
FT SEQUENCE 194 AA; 20463 MW; A19DA4639F811CFE CRC64;  
SQ  
Query Match 3.2%; Score 7; DB 2; Length 194;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 39 AKAILK 45  
DB 146 AKAILK 152

RESULT 25  
Q44997 PRELIMINARY; PRT; 196 AA.  
AC Q44997;

DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE Outer surface protein C (Fragment).  
GN OepC.  
OS Borrelia garinii.  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxId=29519;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SL10;  
RA MEDLINE=95286481; PubMed=7768799;  
RA Thiesen M., Borre M., Mathiesen M.J., Mikkelsen B., Lebech A.M.,  
RA Hansen K.,  
RT "Evolution of the Borrelia burgdorferi outer surface protein OepC.",  
RL J. Bacteriol. 177:3036-3044(1995).  
DR EMBL; X84780; CA959251.1; -  
DR InterPro; IPR01800; Lipoprotein\_6.  
DR Pfam; PF01441; Lipoprotein\_6; 1.  
DR Prodom; PD01149; Lipoprotein\_6; 1.  
FT NON\_TER 1  
FT NON\_TER 196  
FT SEQUENCE 196 AA; 20620 MW; BF1082A759B80821 CRC64;  
SQ  
Query Match 3.2%; Score 7; DB 2; Length 196;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 39 AKAILK 45  
DB 148 AKAILK 154

RESULT 26  
Q9PC51 PRELIMINARY; PRT; 197 AA.  
AC Q9PC51;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE HecI protein.  
GN XF1934.  
OS Xylella fastidiosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
OX Xylella.  
OX NCBI\_TaxId=2371;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=9A5C;  
RX MEDLINE=20365717; PubMed=10910347;  
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,  
RA Alvarenga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,  
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Brites M.R.S.,  
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carter R.,  
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,  
RA Coutinho L.L., Cristofani M., Dias-Neto B., Docena C., El-Dorry H.,  
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,  
RA Fraga J.S., Franca S.C., Franco M.C., Frohne M., Furian L.R.,  
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,  
RA Ho P.L., Hohelsel J.D., Junqueira M.L., Kemper E.B., Kitejima J.P.,  
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,  
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,  
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marukuma A.Y.,  
RA Marques M.V., Martins E.A.L., Martins E.M.F., Marukuma A.Y.,  
RA Menck C.F.M., Miyata E.C., Miyaki C.Y., Montello-Vitorello C.B.,  
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,  
RA Nhani A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
RA de Oliveira M.C., de Oliveira R.C., Palmeri D.A., Paris A.,  
RA Peixoto B.R., Pereira P.G., Pereira H.A.Jr., Pasquero J.B.,  
RA Queglio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
RA de Rosa V.E.Jr., de Sa R.G., Santelli R.V., Sawasaki H.B.,  
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A.Jr.,  
RA da Silveira J.F., Silvestri M.L.Z., Silveira W.J., de Souza A.A.,

OX NCBI\_TaxID=139;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=IP2;  
RX MEDLINE=96296448; PubMed=8709845;  
RA Livey I., Gibbs C.P., Schuster R., Dörner F.;  
RT "Evidence for lateral transfer and recombination in *OspC* variation in  
Lyme disease *Borrelia*.";  
RL Mol. Microbiol. 18:257-269(1995).  
DR EMBL; L42887; AAB36995.1; "-  
DR InterPro; IPR001800; Lipoprotein\_6.  
DR Pfam; PF01441; Lipoprotein\_6; 1.  
DR Prodom; PD001149; Lipoprotein\_6; 1.  
FT NON\_TER 1  
FT NON\_TER 192  
SQ SEQUENCE 192 AA; 20287 MW; 11846F7AC84C7E3D CRC64;  
  
Query Match 3.2%; Score 7; DB 2; Length 192;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 39 AKEALIK 45  
Db 137 AKEALIK 143  
  
RESULT 20  
Q9R7B1 PRELIMINARY; PRT; 192 AA.  
ID Q9R7B1  
AC Q9R7B1;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE Outer surface protein C (Fragment).  
OS *Borrelia burgdorferi* (Lyme disease spirochete).  
OC Bacteria; Spirochaetales; Spirochaetaceae; *Borrelia*.  
OX NCBI\_TaxID=139;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=TETS;  
RX MEDLINE=97478003; PubMed=9336916;  
RA Ras N.M., Postic D., Foretz M., Baranton G.;  
RT "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the  
U.S.A.'";  
RL Int. J. Syst. Bacteriol. 47:1112-1117(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=TETS;  
RA Marti-Ras N., Postic D., Foretz M., Baranton G.;  
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U91798; AAB81895.1; "-  
DR InterPro; IPR001800; Lipoprotein\_6.  
DR Pfam; PF01441; Lipoprotein\_6; 1.  
DR Prodom; PD001149; Lipoprotein\_6; 1.  
FT NON\_TER 1  
FT NON\_TER 192  
SQ SEQUENCE 192 AA; 20297 MW; 6770502A20AA764 CRC64;  
  
Query Match 3.2%; Score 7; DB 2; Length 192;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 39 AKEALIK 45  
Db 141 AKEALIK 147  
  
RESULT 21  
Q9RR53 PRELIMINARY; PRT; 193 AA.  
ID Q9RR53  
AC Q9RR53;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE Outer surface protein C (Fragment).  
GN *OspC*.  
OS *Borrelia burgdorferi* (Lyme disease spirochete).  
OC Bacteria; Spirochaetales; Spirochaetaceae; *Borrelia*.  
OX NCBI\_TaxID=139;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=OCI;  
RX MEDLINE=99091544; PubMed=9872945;  
RA Wang I.N., Dykhuizen D.E., Qiu W., Dunn J.J., Bosler E.M., Luft B.J.;  
RT "Genetic diversity of *ospC* in a local population of *Borrelia*  
*burgdorferi* sensu stricto.";  
RL Genetics 151:15-30(1999).  
DR EMBL; AF029860; AAB8543.1; "-  
DR InterPro; IPR001800; Lipoprotein\_6.  
DR Pfam; PF01441; Lipoprotein\_6; 1.  
DR Prodom; PD001149; Lipoprotein\_6; 1.  
FT NON\_TER 1  
FT NON\_TER 193  
SQ SEQUENCE 193 AA; 20502 MW; 5BFDB5AF8986D1E CRC64;  
  
Query Match 3.2%; Score 7; DB 2; Length 193;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 39 AKEALIK 45  
Db 145 AKEALIK 151  
  
RESULT 22  
Q31115 PRELIMINARY; PRT; 193 AA.  
ID Q31115  
AC Q31115;  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Outer surface protein C (Fragment).  
GN *OspC*.  
OS *Borrelia burgdorferi* (Lyme disease spirochete).  
OC Bacteria; Spirochaetales; Spirochaetaceae; *Borrelia*.  
OX NCBI\_TaxID=139;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=OC3;  
RX MEDLINE=99091544; PubMed=9872945;  
RA Wang I.N., Dykhuizen D.E., Qiu W., Dunn J.J., Bosler E.M., Luft B.J.;  
RT "Genetic diversity of *ospC* in a local population of *Borrelia*  
*burgdorferi* sensu stricto.";  
RL Genetics 151:15-30(1999).  
DR EMBL; AF029862; AAB86545.1; "-  
DR InterPro; IPR001800; Lipoprotein\_6.  
DR Pfam; PF01441; Lipoprotein\_6; 1.  
DR Prodom; PD001149; Lipoprotein\_6; 1.  
FT NON\_TER 1  
FT NON\_TER 193  
SQ SEQUENCE 193 AA; 20596 MW; 018A4CB310475A58 CRC64;  
  
Query Match 3.2%; Score 7; DB 2; Length 193;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 39 AKEALIK 45  
Db 145 AKEALIK 151  
  
RESULT 23  
Q9S517 PRELIMINARY; PRT; 193 AA.  
ID Q9S517  
AC Q9S517;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)

RA Ræ N.M., Postic D., Foretz M., Baranton G.;  
RT "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the  
U.S.A.'?",  
Int. J. Syst. Bacteriol. 47:1112-1117(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=TESTS;  
RA Marti-Rae N., Postic D., Foretz M., Baranton G.;  
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U01797; AAB81894.1; -  
DR InterPro; IPR001800; Lipoprotein\_6.  
DR Pfam; PF01441; Lipoprotein\_6; 1.  
DR Prodom; PD001149; Lipoprotein\_6; 1.  
FT NON\_TER 1  
SQ SEQUENCE 182 AA; 19202 MW; 422146F99A57BF2 CRC64;  
  
Query Match 3.2%; Score 7; DB 2; Length 182;  
Best Local Similarity 100.0%; Pred. No. 53;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 39 AKEAILK 45  
DB 133 AKEAILK 139  
|||||  
P94223 PRELIMINARY; PRT; 191 AA.  
ID P94223;  
AC P94223;  
DT 01-MAY-1997 (TREMBLrel. 03, Created)  
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE Outer surface protein C (Fragment).  
GN OSpC.  
OS Borrelia burgdorferi ( Lyme disease spirochete).  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxID=139;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=M57;  
RX MEDLINE=96296448; PubMed=8709845;  
RA Livey I., Gibbs C.P., Schuster R., Dörner F.;  
RT "Evidence for lateral transfer and recombination in OspC variation in  
Lyme disease Borrelia.",  
RL Mol. Microbiol. 18:257-269(1995).  
DR EMBL; L42879; AAB36988.1; -  
DR InterPro; IPR001800; Lipoprotein\_6.  
DR Pfam; PF01441; Lipoprotein\_6; 1.  
DR Prodom; PD001149; Lipoprotein\_6; 1.  
FT NON\_TER 1  
SQ SEQUENCE 191 AA; 19995 MW; 267D035A6FD7E1C CRC64;  
  
Query Match 3.2%; Score 7; DB 2; Length 191;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 39 AKEAILK 45  
DB 136 AKEAILK 142  
|||||  
P70818 PRELIMINARY; PRT; 191 AA.  
ID P70818;  
AC P70818;  
DT 01-FEB-1997 (TREMBLrel. 02, Created)  
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Outer surface protein (Fragment).  
GN OSpC.  
OS Borrelia burgdorferi ( Lyme disease spirochete).

OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxID=139;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=2-149 CA4;  
RA Probert W.S., Crawford M.R., Cadiz R.B., Lefebvre R.B.;  
RT "Immunization with OspA, but not ospC, provides protection of mice  
challenged with North American isolates of Borrelia burgdorferi.",  
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U81131; AAB06569.1; -  
DR InterPro; IPR001800; Lipoprotein\_6.  
DR Pfam; PF01441; Lipoprotein\_6; 1.  
DR Prodom; PD001149; Lipoprotein\_6; 1.  
FT NON\_TER 1  
SQ SEQUENCE 191 AA; 20126 MW; D2B9B1C02B4DC3C0 CRC64;  
  
Query Match 3.2%; Score 7; DB 2; Length 191;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 39 AKEAILK 45  
DB 136 AKEAILK 142  
|||||  
P94243 PRELIMINARY; PRT; 192 AA.  
ID P94243;  
AC P94243;  
DT 01-MAY-1997 (TREMBLrel. 03, Created)  
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE Outer surface protein C (Fragment).  
GN OSpC.  
OS Borrelia burgdorferi ( Lyme disease spirochete).  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxID=139;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=W;  
RX MEDLINE=96296448; PubMed=8709845;  
RA Livey I., Gibbs C.P., Schuster R., Dörner F.;  
RT "Evidence for lateral transfer and recombination in OspC variation in  
Lyme disease Borrelia.",  
RL Mol. Microbiol. 18:257-269(1995).  
DR EMBL; L42869; AAB37012.1; -  
DR InterPro; IPR001800; Lipoprotein\_6.  
DR Pfam; PF01441; Lipoprotein\_6; 1.  
DR Prodom; PD001149; Lipoprotein\_6; 1.  
FT NON\_TER 1  
SQ SEQUENCE 192 AA; 19981 MW; E9C8A8B304A75388 CRC64;  
  
Query Match 3.2%; Score 7; DB 2; Length 192;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 39 AKEAILK 45  
DB 137 AKEAILK 143  
|||||  
Q9S3P3 PRELIMINARY; PRT; 192 AA.  
ID Q9S3P3;  
AC Q9S3P3;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE Outer surface protein C (Fragment).  
GN OSpC.  
OS Borrelia burgdorferi ( Lyme disease spirochete).  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.

DR Pfam; PF01441; Lipoprotein\_6; 1.  
 DR Pfam; PD001149; Lipoprotein\_6; 1.  
 FT NON\_TER 176 176  
 FT NON\_TER 176 176  
 SQ SEQUENCE 176 AA; 18109 MW; CIA2C029A1FC3DB4 CRC64;

Query Match 3.2%; Score 7; DB 2; Length 176;  
 Best Local Similarity 100.0%; Pred. No. 51;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 AKEATLK 45  
 |||||  
 Db 139 AKEATLK 145

## RESULT 12

Q44999 PRELIMINARY; PRT; 177 AA.

ID 044999  
 AC 044999;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE Outer surface protein C (Fragment).  
 GN OSCP.  
 OS Borrelia burgdorferi ( Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TXGM;  
 RX MEDLINE=95286481; PubMed=7768799;  
 RA Theisen M., Borre M., Mathiesen M.J., Mikkelsen B., Lebech A.M.,  
 RA Hansen K.;  
 RT "Evolution of the Borrelia burgdorferi outer surface protein OSCP";  
 RL J. Bacteriol. 177:3036-3044(1995).  
 DR EMBL; X84783; CA559254.1; -;  
 DR InterPro; IPR001800; Lipoprotein\_6.  
 DR Pfam; PF01441; Lipoprotein\_6; 1.  
 DR Prodom; PD001149; Lipoprotein\_6; 1.  
 FT NON\_TER 177 177  
 FT NON\_TER 177 177  
 SQ SEQUENCE 177 AA; 18800 MW; 156671B9614E7A2D CRC64;

Query Match 3.2%; Score 7; DB 2; Length 177;  
 Best Local Similarity 100.0%; Pred. No. 51;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 AKEATLK 45  
 |||||  
 Db 129 AKEATLK 135

## RESULT 13

Q9REH7 PRELIMINARY; PRT; 178 AA.

ID 09REH7  
 AC 09REH7;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE Outer surface protein C (Fragment).  
 GN OSCP.  
 OS Borrelia burgdorferi ( Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=B31;  
 RX MEDLINE=94075528; PubMed=8253951;  
 RA Theisen M., Frederiksen B., Lebech A.M., Vunst J., Hansen K.;  
 RT "Polymorphism in ospC gene of Borrelia burgdorferi and  
 immunoreactivity of OSCP protein: implications for taxonomy and for  
 use of OSCP protein as a diagnostic antigen";  
 RL J. Clin. Microbiol. 31:2570-2576(1993).

DR EMBL; X73622; CA552001.1; -;  
 DR InterPro; IPR001800; Lipoprotein\_6.  
 DR Pfam; PF01441; Lipoprotein\_6; 1.  
 DR Prodom; PD001149; Lipoprotein\_6; 1.  
 FT NON\_TER 178 178  
 FT NON\_TER 178 178  
 SQ SEQUENCE 178 AA; 18894 MW; D619A64C646F14EB CRC64;

Query Match 3.2%; Score 7; DB 2; Length 178;  
 Best Local Similarity 100.0%; Pred. No. 51;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 AKEATLK 45  
 |||||  
 Db 130 AKEATLK 136

## RESULT 14

Q34123 PRELIMINARY; PRT; 180 AA.

ID 034123  
 AC 034123;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE Outer surface protein C (Fragment).  
 OS Borrelia burgdorferi ( Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TETS;  
 RX MEDLINE=97478003; PubMed=9336916;  
 RA Ras N.M., Poestic D., Foretz M., Baranton G.;  
 RT "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the  
 U.S.A.'?";  
 RL Int. J. Syst. Bacteriol. 47:1112-1117(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TETS;  
 RX Marti-Ras N., Poestic D., Foretz M., Baranton G.;  
 RT Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U91800; AAB81897.1; -;  
 DR InterPro; IPR001800; Lipoprotein\_6.  
 DR Pfam; PF01441; Lipoprotein\_6; 1.  
 DR Prodom; PD001149; Lipoprotein\_6; 1.  
 FT NON\_TER 180 180  
 FT NON\_TER 180 180  
 SQ SEQUENCE 180 AA; 18545 MW; 3D8C7B4A6575A8D4 CRC64;

Query Match 3.2%; Score 7; DB 2; Length 180;  
 Best Local Similarity 100.0%; Pred. No. 52;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 AKEATLK 45  
 |||||  
 Db 132 AKEATLK 138

## RESULT 15

Q9R7B2 PRELIMINARY; PRT; 182 AA.

ID 09R7B2  
 AC 09R7B2;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE Outer surface protein C (Fragment).  
 OS Borrelia burgdorferi ( Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TETS;  
 RX MEDLINE=97478003; PubMed=9336916;

```
AC Q9R7B0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Outer surface protein C (Fragment).
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TERS;
RX MEDLINE=97478003; PubMed=9336916;
RA Ras N.M., Postic D., Foretz M., Baranton G.;
RT "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
   U.S.A.'?";
RL Int. J. Syst. Bacteriol. 47:1112-1117(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TERS;
RX Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U91799; AAB81896.1; -.
DR InterPro; IPR001800; Lipoprotein_6.
DR Pfam; PF01441; Lipoprotein_6; 1.
DR ProDom; PD001149; Lipoprotein_6; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 159 AA; 16861 MW; 5285D422C43DC213 CRC64;

Query Match 3.2%; Score 7; DB 2; Length 159;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 AKKALIK 45
Db 128 AKKALIK 134

RESULT 9
Q9R7A9 PRELIMINARY; PRT; 163 AA.
ID Q9R7A9;
AC Q9R7A9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Outer surface protein C (Fragment).
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TERS;
RX MEDLINE=97478003; PubMed=9336916;
RA Ras N.M., Postic D., Foretz M., Baranton G.;
RT "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
   U.S.A.'?";
RL Int. J. Syst. Bacteriol. 47:1112-1117(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TERS;
RX Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U91801; AAB81898.1; -.
DR InterPro; IPR001800; Lipoprotein_6.
DR Pfam; PF01441; Lipoprotein_6; 1.
DR ProDom; PD001149; Lipoprotein_6; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 163 AA; 17210 MW; 7550141651BD01FF CRC64;

Query Match 3.2%; Score 7; DB 2; Length 163;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 39 AKKALIK 45
Db 138 AKKALIK 144

RESULT 10
Q9R7B4 PRELIMINARY; PRT; 175 AA.
ID Q9R7B4;
AC Q9R7B4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Outer surface protein C (Fragment).
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TERS;
RX MEDLINE=97478003; PubMed=9336916;
RA Ras N.M., Postic D., Foretz M., Baranton G.;
RT "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
   U.S.A.'?";
RL Int. J. Syst. Bacteriol. 47:1112-1117(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TERS;
RX Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U91792; AAB81889.1; -.
DR InterPro; IPR001800; Lipoprotein_6.
DR Pfam; PF01441; Lipoprotein_6; 1.
DR ProDom; PD001149; Lipoprotein_6; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 175 AA; 18573 MW; EC059E7BD3AC3250 CRC64;

Query Match 3.2%; Score 7; DB 2; Length 175;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 AKKALIK 45
Db 144 AKKALIK 150

RESULT 11
Q34119 PRELIMINARY; PRT; 176 AA.
ID Q34119;
AC Q34119;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Outer surface protein C (Fragment).
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TERS;
RX MEDLINE=97478003; PubMed=9336916;
RA Ras N.M., Postic D., Foretz M., Baranton G.;
RT "Borrelia burgdorferi sensu stricto, a bacterial species 'made in the
   U.S.A.'?";
RL Int. J. Syst. Bacteriol. 47:1112-1117(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TERS;
RX Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U91791; AAB81888.1; -.
DR InterPro; IPR001800; Lipoprotein_6.
```



```
AC 08RVW8;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Chalcone isomerase (Fragment).
GN CHI-2.
OS Malus domestica (Apple) (Malus sylvestris).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eucosids I; Rosales; Rosaceae; Maloideae; Malus.
OX NCBI_TaxID=3750;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, EVERESTE X MM106; TISSUE=LEAF;
RA Venise J.-S., Malnoy M., Faize M., Paulin J.-P., Brisset M.-N.;
RT "Modulation of defense responses of Malus during incompatible and
RT compatible interactions with Erwinia amylovora.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL, AF494399; AAM12892.1; -.
KM Isomerase.
FT NON_TER
FT 120
SQ SEQUENCE 120 AA; 12993 MW; 4C0D6146C3887127 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 157 DAEGKAI 163
Db 44 DAEGKAI 50

RESULT 5
ID 08RVW7 PRELIMINARY; PRT; 120 AA.
AC 08RVW7;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Chalcone isomerase (Fragment).
GN CHI-3.
OS Malus domestica (Apple) (Malus sylvestris).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eucosids I; Rosales; Rosaceae; Maloideae; Malus.
OX NCBI_TaxID=3750;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, EVERESTE X MM106; TISSUE=LEAF;
RA Venise J.-S., Malnoy M., Faize M., Paulin J.-P., Brisset M.-N.;
RT "Modulation of defense responses of Malus during incompatible and
RT compatible interactions with Erwinia amylovora.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL, AF494400; AAM12893.1; -.
KM Isomerase.
FT NON_TER
FT 120
SQ SEQUENCE 120 AA; 12954 MW; 8237AF92C38A124C CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 157 DAEGKAI 163
Db 44 DAEGKAI 50

RESULT 6
ID 09SHZ7 PRELIMINARY; PRT; 140 AA.
AC 09SHZ7;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Chalcone isomerase (Fragment).
GN CHI-2.
OS Malus domestica (Apple) (Malus sylvestris).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eucosids I; Rosales; Rosaceae; Maloideae; Malus.
OX NCBI_TaxID=3750;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, EVERESTE X MM106; TISSUE=LEAF;
RA Venise J.-S., Malnoy M., Faize M., Paulin J.-P., Brisset M.-N.;
RT "Modulation of defense responses of Malus during incompatible and
RT compatible interactions with Erwinia amylovora.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL, AF494400; AAM12893.1; -.
KM Isomerase.
FT NON_TER
FT 120
SQ SEQUENCE 120 AA; 12954 MW; 8237AF92C38A124C CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 157 DAEGKAI 163
Db 44 DAEGKAI 50
```

```
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE En/Spm-1-like transposon protein.
GN ATG22080.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., Vanden S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana.";
RL Nature 402:761-768(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL, AC007232; AAD25816.1; -.
SQ SEQUENCE 140 AA; 15338 MW; C0C8D51D4739DAB8 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 155 ADDAEGK 161
Db 46 ADDAEGK 52

RESULT 7
ID 015401 PRELIMINARY; PRT; 157 AA.
AC 015401;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JAN-1998 (TREMBlrel. 05, Last annotation update)
DE RNA polymerase transcriptional regulation mediator.
GN H-MED6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kim Y.-J.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL, U78082; AAB84363.1; -.
SQ SEQUENCE 157 AA; 18101 MW; 2894BEF4D2B9BF9F CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 170 LNSGSVL 176
Db 24 LNSGSVL 30

RESULT 8
ID 09R7B0 PRELIMINARY; PRT; 159 AA.
AC 09R7B0;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Chalcone isomerase (Fragment).
GN CHI-2.
OS Malus domestica (Apple) (Malus sylvestris).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eucosids I; Rosales; Rosaceae; Maloideae; Malus.
OX NCBI_TaxID=3750;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, EVERESTE X MM106; TISSUE=LEAF;
RA Venise J.-S., Malnoy M., Faize M., Paulin J.-P., Brisset M.-N.;
RT "Modulation of defense responses of Malus during incompatible and
RT compatible interactions with Erwinia amylovora.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL, AF494400; AAM12893.1; -.
KM Isomerase.
FT NON_TER
FT 120
SQ SEQUENCE 120 AA; 12954 MW; 8237AF92C38A124C CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 170 LNSGSVL 176
Db 24 LNSGSVL 30
```

966	6	2.8	468	2	Q8RKY2	Q8RKY2 streptococc
967	6	2.8	468	10	Q9ZVA7	Q9ZVA7 aradidopsis
968	6	2.8	468	16	Q8XYM9	Q8XYM9 anabaena sp
969	6	2.8	469	2	O66233	O66233 escherichia
970	6	2.8	470	16	Q8YX86	Q8YX86 anabaena sp
971	6	2.8	472	10	Q8R2S4	Q8R2S4 oryza sativ
972	6	2.8	473	2	P97002	P97002 streptococc
973	6	2.8	473	10	Q8W250	Q8W250 oryza sativ
974	6	2.8	473	17	Q9YEU4	Q9YEU4 aeropyrum p
975	6	2.8	475	2	Q8EYV0	Q8EYV0 klebsiella
976	6	2.8	475	16	Q8X7X7	Q8X7X7 escherichia
977	6	2.8	476	3	O59677	O59677 schistosach
978	6	2.8	476	16	Q9XP20	Q9XP20 vibrio chol
979	6	2.8	477	3	Q9UM05	Q9UM05 clavispora
980	6	2.8	478	5	O17392	O17392 caenorhabdi
981	6	2.8	478	16	Q8ZBP2	Q8ZBP2 yerania pe
982	6	2.8	478	16	Q8XFR3	Q8XFR3 salmonella
983	6	2.8	479	16	Q8ZMF5	Q8ZMF5 salmonella
984	6	2.8	479	16	Q8Z470	Q8Z470 salmonella
985	6	2.8	481	16	Q8YHK7	Q8YHK7 brucella me
986	6	2.8	482	2	O9AJE2	O9AJE2 kitasatospo
987	6	2.8	483	16	Q8R8Q1	Q8R8Q1 thermoaer
988	6	2.8	484	5	Q9NVLS	Q9NVLS homo sapien
989	6	2.8	484	5	O9BKX1	O9BKX1 naegleria f
990	6	2.8	486	16	Q8ZE12	Q8ZE12 yerania pe
991	6	2.8	486	16	Q8XH03	Q8XH03 salmonella
992	6	2.8	486	16	Q8XC06	Q8XC06 escherichia
993	6	2.8	486	16	Q9X8J9	Q9X8J9 streptomyc
994	6	2.8	487	10	Q9SX11	Q9SX11 arabidopsis
995	6	2.8	488	16	Q8XXW1	Q8XXW1 ralslonia s
996	6	2.8	489	5	O9V6Q7	O9V6Q7 dirocephila
997	6	2.8	489	16	Q8YPA1	Q8YPA1 anabaena sp
998	6	2.8	490	4	O75125	O75125 homo sapien
999	6	2.8	490	11	O920K3	O920K3 rattus norv
1000	6	2.8	492	4	Q9ET73	Q9ET73 homo sapien

## ALIGNMENTS

RESULT 1

Q45178 ID Q45178 PRELIMINARY; PRT; 210 AA.

AC Q45178; DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)

DE Outer surface protein C.

OS Borrelia garinii.

OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.

NCBI\_TaxID=29519;

CC NCB1\_TaxID=29519;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=HT64;

RX MEDLINE=96025162; PubMed=7494039;

RA Fukunaga M., Hamase A.;

RT "Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu lato isolates from Japan."

RL J. Clin. Microbiol. 33:2415-2420 (1995).

DR EMBL; D49378; BAA08376.1; -

DR InterPro; IPR001800; Lipoprotein\_6.

DR Pfam; PF01441; Lipoprotein\_6; 1.

DR ProDom; PDD01149; Lipoprotein\_6; 1.

SQ SEQUENCE 210 AA; 21973 MW; EED09CF942A599E0 CRC64;

Query Match 3.7%; Score 8; DB 2; Length 210;

Best Local Similarity 100.0%; Pred. No. 5.1;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 LAKEATLK 45

DB 154 LAKEATLK 161

RESULT 2

Q47616 ID Q47616 PRELIMINARY; PRT; 120 AA.

AC Q47616; DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE ORF120.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

CC Escherichia.

NCBI\_TaxID=562;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=K-12;

RA Miyamoto K., Inokuchi H.;

RT "Nucleotide sequence of 5' flanking region of the ribosomal RNA gene (rrnH) in E. coli."

RL Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.

DR EMBL; D15061; BAA03660.1; -

DR InterPro; IPR003439; ABC\_transport.

DR Pfam; PF00005; ABC\_tran; 1.

SQ SEQUENCE 120 AA; 13232 MW; B6971E2C875A6E0C CRC64;

Query Match 3.2%; Score 7; DB 2; Length 120;

Best Local Similarity 100.0%; Pred. No. 35;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 GSVLVDG 179

DB 60 GSVLVDG 66

RESULT 3

Q8RVM9 ID Q8RVM9 PRELIMINARY; PRT; 120 AA.

AC Q8RVM9; DT 01-JUN-2002 (TREMBlrel. 21, Created)

DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Chalcone isomerase (Fragment).

GN CHI-1.

OS Malus domestica (Apple) (Malus sylvestris).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC Eucosids I; Rosales; Rosaceae; Maloideae; Malus.

NCBI\_TaxID=3750;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. EVERESTE X MM106; TISSUE=LEAF;

RA Venisse J.-S., Malnoy M., Patze M., Paulin J.-P., Brieset M.-N.;

RT "Modulation of defense responses of Malus during incompatible and RT compatible interactions with Erwinia amylovora."

RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF494398; AAM12891.1; -

KW Isomerase.

FT NON\_TER 1 1

FT NON\_TER 120 120

SQ SEQUENCE 120 AA; 12978 MW; 33F511A25B4AB8C3 CRC64;

Query Match 3.2%; Score 7; DB 10; Length 120;

Best Local Similarity 100.0%; Pred. No. 35;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 DAEGKAI 163

DB 44 DAEGKAI 50

RESULT 4

Q8RVM8 ID Q8RVM8 PRELIMINARY; PRT; 120 AA.

820	6	2.8	396	2	Q9AMF5	Q9amf5 bruceella ab	893	6	2.8	441	12	Q9YTN4	Q9ytn4 ateline her
821	6	2.8	396	16	Q9YDL6	Q9ydl6 bruceella me	894	6	2.8	441	17	Q979Z3	Q979z3 thermoplas
822	6	2.8	397	10	Q94UD9	Q94jd9 arabidopsis	895	6	2.8	442	5	Q9VSM2	Q9vsm2 drosophila
823	6	2.8	398	16	Q8XJN2	Q8xjn2 clostridium	896	6	2.8	442	10	Q9XIO7	Q9xio7 arabidopsis
824	6	2.8	399	5	Q8XOA9	Q8xoa9 neurospora	897	6	2.8	442	10	Q94A13	Q94a13 arabidopsis
825	6	2.8	399	5	Q16765	Q16765 caenorhabdi	898	6	2.8	442	10	Q8YV77	Q8yv77 arabidopsis
826	6	2.8	400	2	Q10788	Q10788 rhodococcus	899	6	2.8	442	10	Q8RU74	Q8ru74 lycopersico
827	6	2.8	403	3	Q42963	Q42963 schizosacch	900	6	2.8	442	11	Q8RID4	Q8rid4 mus musculu
828	6	2.8	403	5	Q8T2W3	Q8t2w3 trypanosoma	901	6	2.8	443	3	Q9C1R4	Q9c1r4 leontium e
829	6	2.8	404	15	Q9K1R0	Q9k1r0 neisseria m	902	6	2.8	444	2	Q69240	Q69240 bradyrhizob
830	6	2.8	404	17	Q97A08	Q97a8 thermoplas	903	6	2.8	444	3	Q96VU2	Q96vu2 leontium e
831	6	2.8	406	5	Q9XTR1	Q9xtr1 caenorhabdi	904	6	2.8	444	11	Q99KN1	Q99kn1 mus musculu
832	6	2.8	406	10	Q9FP30	Q9ff30 arabidopsis	905	6	2.8	444	16	Q8UPE7	Q8upe7 agrobacteri
833	6	2.8	406	11	Q9C203	Q9cz03 mus musculu	906	6	2.8	446	9	Q8W702	Q8w702 cyanophaga
834	6	2.8	406	13	Q9W6N3	Q9w6n3 gallus gall	907	6	2.8	446	16	Q98A08	Q98a08 rhizobium 1
835	6	2.8	407	16	Q9ZDS9	Q9zds9 rickettsia	908	6	2.8	447	2	Q9AHR9	Q9ahr9 acetogenium
836	6	2.8	407	16	Q8UIZ1	Q8uiz1 agrobacteri	909	6	2.8	447	16	Q9HME2	Q9hme2 pseudomonas
837	6	2.8	407	17	Q9YAS1	Q9yaa1 aeropyrum p	910	6	2.8	447	16	Q97G47	Q97g47 clostridium
838	6	2.8	408	6	Q9SK13	Q9sk13 macaca fasc	911	6	2.8	450	12	Q89568	Q89568 rabies viru
839	6	2.8	408	16	Q98BD4	Q98bd4 rhizobium 1	912	6	2.8	450	12	Q89568	Q89568 rabies viru
840	6	2.8	408	16	Q92IU0	Q92iu0 rickettsia	913	6	2.8	450	12	Q86417	Q86417 rabies viru
841	6	2.8	408	17	Q97X39	Q97x39 sulfolobus	914	6	2.8	450	12	Q86421	Q86421 rabies viru
842	6	2.8	409	16	Q9ABW3	Q9abw3 caulobacter	915	6	2.8	450	12	Q86424	Q86424 rabies viru
843	6	2.8	409	16	Q8XJ75	Q8xj15 clostridium	916	6	2.8	450	12	Q86429	Q86429 rabies viru
844	6	2.8	409	16	Q07238	Q07238 mycobacteri	917	6	2.8	450	12	Q86434	Q86434 rabies viru
845	6	2.8	409	16	Q92MT6	Q92mt6 rhizobium m	918	6	2.8	450	12	Q86441	Q86441 rabies viru
846	6	2.8	410	10	Q9XEB6	Q9xeb6 sorghum bic	919	6	2.8	450	12	Q86443	Q86443 rabies viru
847	6	2.8	412	17	Q96XP8	Q96xp8 sulfolobus	920	6	2.8	450	12	Q86444	Q86444 rabies viru
848	6	2.8	412	17	Q9V295	Q9v295 pyrococcus	921	6	2.8	450	12	Q86445	Q86445 rabies viru
849	6	2.8	413	16	P96860	P96860 mycobacteri	922	6	2.8	450	12	Q86446	Q86446 rabies viru
850	6	2.8	413	16	Q9S2M3	Q9s2m3 streptomyce	923	6	2.8	450	12	Q86447	Q86447 rabies viru
851	6	2.8	414	10	Q9M8W9	Q9m8w9 arabidopsis	924	6	2.8	450	12	Q86479	Q86479 rabies viru
852	6	2.8	415	16	Q9KRP6	Q9krc6 vibrio chol	925	6	2.8	450	12	Q86480	Q86480 rabies viru
853	6	2.8	416	3	Q9HGJ7	Q9hgj7 saccharomyc	926	6	2.8	450	12	Q918X2	Q918x2 rabies viru
854	6	2.8	416	3	Q9HGJ6	Q9hgj6 saccharomyc	927	6	2.8	450	12	Q918X0	Q918x0 rabies viru
855	6	2.8	416	10	Q9M0N4	Q9mon4 arabidopsis	928	6	2.8	450	12	Q91SV0	Q91sv0 rabies viru
856	6	2.8	418	10	Q9CT05	Q9ct05 arabidopsis	929	6	2.8	450	12	Q91SV0	Q91sv0 rabies viru
857	6	2.8	419	10	Q9SMQ9	Q9smq9 euglena gra	930	6	2.8	450	12	Q91SU4	Q91su4 rabies viru
858	6	2.8	420	16	Q9RRQ7	Q9rrq7 deinococcus	931	6	2.8	450	12	Q91SU3	Q91su3 rabies viru
859	6	2.8	422	4	Q16773	Q16773 homo sapien	932	6	2.8	450	12	Q91845	Q91845 rabies viru
860	6	2.8	423	10	Q9L186	Q9l186 arabidopsis	933	6	2.8	450	12	Q80Y66	Q80y66 rabies viru
861	6	2.8	424	5	Q9SSA2	Q9ssa2 drosophila	934	6	2.8	451	5	Q21520	Q21520 caenorhabdi
862	6	2.8	424	5	Q9VZ08	Q9vz08 drosophila	935	6	2.8	451	12	Q66534	Q66534 european ba
863	6	2.8	424	16	P72922	P72922 synecchocyst	936	6	2.8	451	12	Q66535	Q66535 european ba
864	6	2.8	425	2	Q86101	Q86101 arthrobacte	937	6	2.8	451	16	Q986L8	Q986l8 rhizobium 1
865	6	2.8	425	2	Q93NH4	Q93nh4 arthrobacte	938	6	2.8	451	16	Q8RECT3	Q8rect3 thermoaer
866	6	2.8	425	10	Q9SUT6	Q9sje6 arabidopsis	939	6	2.8	452	2	Q9A0S4	Q9a0s4 burkholderi
867	6	2.8	426	16	Q9KUD8	Q9ky0 cyza sativ	940	6	2.8	452	2	Q52922	Q52922 clostridium
868	6	2.8	426	16	Q9JVE6	Q9jve6 neisseria m	941	6	2.8	452	16	Q9L253	Q9l253 streptomyce
869	6	2.8	427	2	Q9FBJ7	Q9fbj7 streptomyce	942	6	2.8	452	17	Q97W75	Q97w75 pseudomonas
870	6	2.8	427	2	Q9FF71	Q9ff71 pseudomonas	943	6	2.8	452	17	Q9V246	Q9v246 pyrococcus
871	6	2.8	427	10	Q82490	Q82490 arabidopsis	944	6	2.8	453	17	Q975B2	Q975b2 sulfolobus
872	6	2.8	427	12	Q65792	Q65792 mucosal dis	945	6	2.8	454	4	Q9HAB6	Q9hab6 homo sapien
873	6	2.8	427	12	Q65792	Q65792 mucosal dis	946	6	2.8	454	16	Q9RVK1	Q9rvk1 deinococcus
874	6	2.8	427	16	Q8YP40	Q8yp40 anabaena sp	947	6	2.8	455	16	Q970E1	Q97je1 clostridium
875	6	2.8	428	10	Q9M2X8	Q9m2x8 arabidopsis	948	6	2.8	456	3	Q06340	Q06340 saccharomyc
876	6	2.8	429	16	Q9PN14	Q9pn14 campylobact	949	6	2.8	457	5	Q9N2M5	Q9n2m5 caenorhabdi
877	6	2.8	429	16	Q98CB6	Q98ce6 rhizobium 1	950	6	2.8	457	5	Q17463	Q17463 caenorhabdi
878	6	2.8	429	16	Q8YD54	Q8yds4 bruceella me	951	6	2.8	457	5	Q960D1	Q960d1 drosophila
879	6	2.8	431	5	Q966H7	Q966h7 caenorhabdi	952	6	2.8	458	5	Q95X02	Q95x02 naegleria f
880	6	2.8	431	10	Q8S4Q4	Q8s4q4 glycine max	953	6	2.8	460	3	Q02321	Q02321 phanerocha
881	6	2.8	431	16	Q9KSK3	Q9ksk3 vibrio chol	954	6	2.8	460	6	Q95JX8	Q95jx8 macaca fasc
882	6	2.8	432	5	Q8T2S6	Q8t2s6 dictyosteli	955	6	2.8	463	2	Q91AX4	Q91ax4 streptococc
883	6	2.8	433	11	Q8VEG7	Q8veg7 mus musculu	956	6	2.8	463	2	Q93AB6	Q93ab6 yersinia pe
884	6	2.8	434	16	Q9WY03	Q9wy03 thermotoga	957	6	2.8	464	13	Q91913	Q91913 xenopus lae
885	6	2.8	435	2	Q9FCN8	Q9fcn8 rhizobium 1	958	6	2.8	464	13	Q91693	Q91693 xenopus lae
886	6	2.8	435	2	Q00564	Q00564 rhizobium 1	959	6	2.8	465	16	Q9AEN3	Q9aen3 oryza sativ
887	6	2.8	435	9	Q8SD69	Q8sd69 pseudomonas	960	6	2.8	465	16	Q92083	Q92q83 rhizobium m
888	6	2.8	435	10	Q9XEG9	Q9xeg9 sorghum bic	961	6	2.8	466	16	Q9BHE2	Q9bhe2 rhizobium 1
889	6	2.8	437	15	Q8ZH13	Q8zh13 yersinia pe	962	6	2.8	467	4	Q9BXS0	Q9bxs0 homo sapien
890	6	2.8	438	17	Q8U0Q3	Q8u0q3 pyrococcus	963	6	2.8	467	13	P79947	P79947 xenopus lae
891	6	2.8	439	2	Q85833	Q85833 sphingomona	964	6	2.8	467	16	Q92P20	Q92p20 rhizobium m
892	6	2.8	441	10	Q9AYD7	Q9ayd7 oryza sativ	965	6	2.8	468	2	Q9ZJ01	Q9zj01 streptococc

674	6	2.8	340	10	Q9AUM5	Q9aun5 oryza sativ	747	6	2.8	364	16	Q8YLR6	Q8y1r6 anabaena sp
675	6	2.8	340	16	Q8UTR5	Q8ut5 agrobacteri	748	6	2.8	365	5	Q211L5	Q211l5 caenorhabdi
676	6	2.8	341	2	Q9AGY8	Q9ag8 agrobacteri	749	6	2.8	366	17	Q9V1D0	Q9v1d0 pyrococcus
677	6	2.8	341	16	Q97M60	Q97m6 clostridium	750	6	2.8	367	16	Q8UBA5	Q8ub5 agrobacteri
678	6	2.8	342	5	Q9XTB6	Q9xtb6 caenorhabdi	751	6	2.8	367	17	Q8UTW5	Q8utw5 methanopyru
679	6	2.8	342	10	Q9AUM7	Q9aum7 oryza sativ	752	6	2.8	368	16	Q8YEB8	Q8yeb8 bruceella me
680	6	2.8	342	16	Q9A4F7	Q9a4f7 caulobacter	753	6	2.8	370	5	Q9NAH3	Q9nah3 caenorhabdi
681	6	2.8	342	17	Q9V192	Q9v192 pyrococcus	754	6	2.8	370	16	Q9A434	Q9a434 caulobacter
682	6	2.8	343	5	Q9XMA4	Q9xma4 caenorhabdi	755	6	2.8	370	16	Q9A217	Q9a217 caulobacter
683	6	2.8	343	10	Q9M1P6	Q9m1p6 arabadopsis	756	6	2.8	370	16	Q92BD6	Q92bd6 listeria in
684	6	2.8	343	17	Q97M64	Q97m64 sulfolobus	757	6	2.8	370	16	Q8TV62	Q8tv62 listeria mo
685	6	2.8	344	10	Q8M3E1	Q8m3e1 oryza sativ	758	6	2.8	370	16	Q8UC98	Q8uc98 agrobacteri
686	6	2.8	344	17	Q9UYO6	Q9uy6 pyrococcus	759	6	2.8	371	2	Q8RME8	Q8rme8 amycolatops
687	6	2.8	345	12	Q68543	Q68543 horseradish	760	6	2.8	371	16	Q98N04	Q98n04 rhizobium 1
688	6	2.8	347	10	Q68V53	Q68v53 rhodomonas	761	6	2.8	371	16	Q982G3	Q982g3 rhizobium 1
689	6	2.8	347	16	Q9PHU0	Q9ph10 xyliella faa	762	6	2.8	371	16	Q9S227	Q9s227 streptomyce
690	6	2.8	347	16	Q9ZUJ9	Q9zu19 rhizobium m	763	6	2.8	371	17	Q8THB3	Q8thb3 methanosarc
691	6	2.8	348	16	Q9AK41	Q9ak41 streptomyce	764	6	2.8	372	5	Q9V588	Q9v588 drosophila
692	6	2.8	350	5	Q19582	Q19582 caenorhabdi	765	6	2.8	372	17	Q8TX20	Q8tx20 methanopyru
693	6	2.8	350	13	Q9PVV1	Q9pvv1 letheion	766	6	2.8	372	5	Q9W1X2	Q9w1x2 drosophila
694	6	2.8	351	2	Q87241	Q87241 lactococcus	767	6	2.8	373	11	Q9R0P2	Q9rop2 mus musculu
695	6	2.8	351	12	Q82676	Q82676 indian cass	768	6	2.8	373	17	Q8ZVN9	Q8zvn9 pyrobaculu
696	6	2.8	351	17	Q9H187	Q9h187 thermoplasm	769	6	2.8	374	2	Q05107	Q05107 ectochiorho
697	6	2.8	351	17	Q97CG9	Q97cg9 thermoplasm	770	6	2.8	374	10	Q9LHW8	Q9lhw8 oryza sativ
698	6	2.8	352	5	Q9V5F1	Q9v5f1 drosophila	771	6	2.8	375	2	Q93P55	Q93p55 myxococcus
699	6	2.8	352	12	Q8U255	Q8u255 indian cass	772	6	2.8	375	16	Q92V13	Q92v13 rhizobium m
700	6	2.8	353	16	Q98D44	Q98d44 rhizobium 1	773	6	2.8	377	3	Q9HGK0	Q9hgk0 saccharomyc
701	6	2.8	353	16	Q98BM1	Q98bm1 rhizobium 1	774	6	2.8	377	3	Q9HGJ9	Q9hgj9 saccharomyc
702	6	2.8	353	16	Q8UD21	Q8ud21 agrobacteri	775	6	2.8	377	3	Q9HGJ8	Q9hgj8 saccharomyc
703	6	2.8	354	10	Q9M5P3	Q9m5p3 arabadopsis	776	6	2.8	377	10	Q94EJ2	Q94ej2 arabadopsis
704	6	2.8	354	10	Q93Y53	Q93y53 arabadopsis	777	6	2.8	377	11	Q9CRR7	Q9crr7 mus musculu
705	6	2.8	355	5	Q77112	Q77112 automeria i	778	6	2.8	377	16	Q92V91	Q92v91 rhizobium m
706	6	2.8	356	5	Q966F9	Q966f9 caenorhabdi	779	6	2.8	377	17	Q97YD5	Q97yds sulfolobus
707	6	2.8	356	12	Q918D4	Q918d4 chilli leaf	780	6	2.8	378	5	Q8T8S2	Q8t8s2 drosophila
708	6	2.8	356	16	Q995H4	Q995h4 staphylococ	781	6	2.8	378	10	Q9FXK0	Q9fxk0 arabadopsis
709	6	2.8	356	17	Q9HK04	Q9hk04 thermoplasm	782	6	2.8	378	10	Q9M0E6	Q9m0e6 arabadopsis
710	6	2.8	357	12	Q38544	Q38544 tomato yell	783	6	2.8	378	10	Q8R2C4	Q8rcz4 oryza sativ
711	6	2.8	358	4	Q9BRP1	Q9brp1 homo sapien	784	6	2.8	378	12	Q99H03	Q99h03 helicoverpa
712	6	2.8	358	10	Q9MAN1	Q9man1 arabadopsis	785	6	2.8	378	12	Q8V5U6	Q8v5u6 helicoverpa
713	6	2.8	358	12	Q9WR17	Q9wr17 african cas	786	6	2.8	378	16	Q97E48	Q97e48 clostridium
714	6	2.8	358	12	Q9JEA9	Q9jea9 casabaava gem	787	6	2.8	379	2	Q45467	Q45467 bacillus sp
715	6	2.8	358	12	Q9JEA2	Q9jea2 casabaava gem	788	6	2.8	379	13	Q918Y8	Q918y8 brachydanio
716	6	2.8	358	12	Q8QUW0	Q8quw0 african cas	789	6	2.8	379	16	Q9A044	Q9a044 streptococc
717	6	2.8	359	5	Q8T130	Q8t130 dictyosteli	790	6	2.8	379	16	Q97HK2	Q97hk2 clostridium
718	6	2.8	359	12	Q88942	Q88942 tomato yell	791	6	2.8	380	5	Q8W011	Q8w011 retraymena
719	6	2.8	359	12	Q9Y127	Q9y127 tomato yell	792	6	2.8	380	10	Q43038	Q43038 populus bal
720	6	2.8	359	12	Q9Y2V4	Q9y2v4 tomato yell	793	6	2.8	380	16	Q8V4Z6	Q8v4z6 bruceella me
721	6	2.8	359	12	Q9Y2V2	Q9y2v2 tomato yell	794	6	2.8	380	17	Q97245	Q97245 sulfolobus
722	6	2.8	359	12	Q9YUX7	Q9yux7 tomato yell	795	6	2.8	381	10	Q9A5C1	Q9a5c1 oryza sativ
723	6	2.8	359	12	Q8V380	Q8v380 east africa	796	6	2.8	381	16	Q98HU8	Q98hu8 rhizobium 1
724	6	2.8	359	16	Q54126	Q54126 streptomyce	797	6	2.8	381	17	Q97J38	Q97j38 sulfolobus
725	6	2.8	360	12	Q65317	Q65317 egeratum ye	798	6	2.8	382	5	Q16794	Q16794 caenorhabdi
726	6	2.8	360	12	Q9QC33	Q9qc33 tomato leaf	799	6	2.8	382	5	Q18391	Q18391 caenorhabdi
727	6	2.8	360	17	Q97W70	Q97w70 sulfolobus	800	6	2.8	383	4	Q9NTX9	Q9ntx9 homo sapien
728	6	2.8	361	12	Q9WH24	Q9wh24 indian mung	801	6	2.8	384	10	Q9SEF78	Q9sef78 arabadopsis
729	6	2.8	361	12	Q9YPS2	Q9yps2 vigna mungo	802	6	2.8	384	10	Q9LNU3	Q9lnu3 arabadopsis
730	6	2.8	361	12	Q91597	Q91597 mungbean ye	803	6	2.8	385	10	Q94G10	Q94g10 oryza sativ
731	6	2.8	361	12	Q8V980	Q8v980 soybean cri	804	6	2.8	385	10	Q947T6	Q947t6 oryza sativ
732	6	2.8	361	17	Q59390	Q59390 pyrococcus	805	6	2.8	386	16	Q9K3U6	Q9k3u6 bacillus ha
733	6	2.8	362	11	Q9JHQ1	Q9jhq1 tabacus norv	806	6	2.8	386	16	Q8YB28	Q8yb28 bruceella me
734	6	2.8	362	12	Q9IGX9	Q9igx9 tabacco lea	807	6	2.8	387	10	Q9LNB0	Q9lnb0 oryza sativ
735	6	2.8	362	12	Q99DR7	Q99dr7 vigna mungo	808	6	2.8	387	17	Q8U046	Q8u046 pyrococcus
736	6	2.8	362	12	Q992L5	Q992l5 tomato leaf	809	6	2.8	388	17	Q8ZVR9	Q8zvr9 pyrobaculu
737	6	2.8	362	12	Q915C1	Q915c1 indian mung	810	6	2.8	389	16	Q98AK0	Q98ak0 rhizobium 1
738	6	2.8	362	12	Q917N1	Q917n1 indian mung	811	6	2.8	389	16	Q8KHX0	Q8kxh0 clostridium
739	6	2.8	362	12	Q913B9	Q913b9 indian mung	812	6	2.8	390	17	Q8TQ17	Q8tq17 methanosarc
740	6	2.8	362	12	Q910Y7	Q910y7 soybean yell	813	6	2.8	391	16	Q9ABZ1	Q9abz1 caulobacter
741	6	2.8	362	12	Q8V011	Q8v011 bolllyhook 1	814	6	2.8	392	16	Q9JWS8	Q9jws8 neiseeria m
742	6	2.8	362	12	Q8UYT6	Q8uyt6 soybean yell	815	6	2.8	392	13	Q97K03	Q97k03 clostridium
743	6	2.8	362	17	Q9HSE6	Q9hse6 halobacteri	816	6	2.8	394	16	Q98SK7	Q98sk7 rana dybows
744	6	2.8	363	10	Q9LPZ0	Q9lpz0 arabadopsis	817	6	2.8	395	10	Q9FGI5	Q9fgi5 arabadopsis
745	6	2.8	363	16	Q9RJL9	Q9rjl9 streptomyce	818	6	2.8	395	10	Q8S1Y2	Q8s1y2 oryza sativ
746	6	2.8	364	11	Q9D1M3	Q9dlm3 mus musculu	819	6	2.8	395	17	Q58134	Q58134 pyrococcus

528	6	2.8	287	4	Q9P0P8	Q9P0F8 homo sapien	601	6	2.8	307	17	059053	059053 pyrococcus
529	6	2.8	288	5	002361	002361 caenorhabdi	602	6	2.8	308	4	Q9BU10	Q9BU10 homo sapien
530	6	2.8	288	17	Q976F5	Q976F5 sulfolobus	603	6	2.8	308	11	Q8VD14	Q8VD14 mus musculu
531	6	2.8	289	2	Q9XBK8	Q9XBK8 bacillus ce	604	6	2.8	308	12	Q89369	Q89369 parameciu
532	6	2.8	289	12	Q9WH91	Q9WH91 kaposi's sa	605	6	2.8	308	16	Q9KCA5	Q9KCA5 bacillus ha
533	6	2.8	289	12	Q9WH92	Q9WH92 kaposi's sa	606	6	2.8	309	3	Q9USL1	Q9USL1 schizosacch
534	6	2.8	289	12	Q9WHB6	Q9WHB6 kaposi's sa	607	6	2.8	309	10	Q9FKM1	Q9FKM1 arabiidopsi
535	6	2.8	289	12	Q9WHB7	Q9WHB7 kaposi's sa	608	6	2.8	309	10	Q93J21	Q93J21 arabiidopsi
536	6	2.8	289	12	Q9WHC5	Q9WHC5 kaposi's sa	609	6	2.8	310	16	P96562	P96562 amycolatops
537	6	2.8	289	12	Q9WH10	Q9WH10 kaposi's sa	610	6	2.8	310	16	Q9HTL4	Q9HTL4 pseudomonas
538	6	2.8	289	12	Q9Q0G6	Q9Q0G6 kaposi's sa	611	6	2.8	310	16	Q9XBR9	Q9XBR9 streptomyce
539	6	2.8	289	12	Q9DSR4	Q9DSR4 kaposi's sa	612	6	2.8	311	5	Q95YCO	Q95YCO caenorhabdi
540	6	2.8	289	12	Q9DSR5	Q9DSR5 kaposi's sa	613	6	2.8	312	12	Q90Z38	Q90Z38 porcine cir
541	6	2.8	289	12	Q9DSR7	Q9DSR7 kaposi's sa	614	6	2.8	312	12	P89032	P89032 porcine cir
542	6	2.8	289	12	Q9DSR6	Q9DSR6 kaposi's sa	615	6	2.8	312	16	Q9HWK3	Q9HWK3 pseudomonas
543	6	2.8	289	12	Q9DSR3	Q9DSR3 kaposi's sa	616	6	2.8	313	12	Q9DW12	Q9DW12 pluteella xy
544	6	2.8	289	12	Q9DSR9	Q9DSR9 kaposi's sa	617	6	2.8	313	16	Q929F3	Q929F3 listeria in
545	6	2.8	289	12	Q9DSR8	Q9DSR8 kaposi's sa	618	6	2.8	313	16	Q8Y556	Q8Y556 listeria mo
546	6	2.8	289	12	Q9DSR7	Q9DSR7 kaposi's sa	619	6	2.8	314	11	Q8VGL1	Q8VGL1 mus musculu
547	6	2.8	289	12	Q9DSC6	Q9DSC6 kaposi's sa	620	6	2.8	314	11	Q8YKAB	Q8YKAB anabaena sp
548	6	2.8	289	12	Q9DSC3	Q9DSC3 kaposi's sa	621	6	2.8	314	16	Q65535	Q65535 aquifex aeo
549	6	2.8	289	12	Q9DSC1	Q9DSC1 kaposi's sa	622	6	2.8	316	4	Q9NTA3	Q9NTA3 homo sapien
550	6	2.8	289	12	Q9DSC0	Q9DSC0 kaposi's sa	623	6	2.8	316	5	Q9GPE6	Q9GPE6 dictyosteli
551	6	2.8	289	12	Q9DXH2	Q9DXH2 beak and fe	624	6	2.8	317	8	Q9BAP6	Q9BAP6 candida alb
552	6	2.8	289	12	Q9DXH0	Q9DXH0 beak and fe	625	6	2.8	317	13	Q9PST4	Q9PST4 ictalurus p
553	6	2.8	289	16	Q69838	Q69838 streptomyce	626	6	2.8	318	17	Q8TRU1	Q8TRU1 methanosarc
554	6	2.8	290	4	Q9BUV3	Q9BUV3 homo sapien	627	6	2.8	319	2	Q87367	Q87367 etaphylococ
555	6	2.8	290	11	Q921C9	Q921C9 cricetulus	628	6	2.8	319	3	P78869	P78869 schizosacch
556	6	2.8	290	16	Q8U679	Q8U679 agrobacteri	629	6	2.8	319	8	Q34834	Q34834 kluyveromyc
557	6	2.8	291	2	Q24740	Q24740 streptomyce	630	6	2.8	319	17	Q8ZWS2	Q8ZWS2 pyrobaculum
558	6	2.8	291	16	Q9CIB5	Q9CIB5 lactococcus	631	6	2.8	320	4	Q9BT76	Q9BT76 homo sapien
559	6	2.8	292	2	Q9Z131	Q9Z131 enterococcu	632	6	2.8	320	5	Q9VF44	Q9VF44 drosophila
560	6	2.8	293	16	Q8X8J1	Q8X8J1 escherichia	633	6	2.8	320	12	Q91DM6	Q91DM6 soybean yel
561	6	2.8	294	2	Q8VPN8	Q8VPN8 micrococcus	634	6	2.8	320	12	Q8VJMI	Q8VJMI swinepox vi
562	6	2.8	294	3	Q12219	Q12219 saccharomyc	635	6	2.8	321	10	Q41412	Q41412 solanum tub
563	6	2.8	294	9	Q9AZV4	Q9AZV4 bacterioph	636	6	2.8	321	10	Q41413	Q41413 solanum tub
564	6	2.8	294	9	Q9AZP7	Q9AZP7 bacterioph	637	6	2.8	321	10	Q41414	Q41414 solanum tub
565	6	2.8	294	16	Q9CFP6	Q9CFP6 lactococcus	638	6	2.8	322	10	Q41415	Q41415 solanum tub
566	6	2.8	295	5	Q9XZP5	Q9XZP5 brugia mala	639	6	2.8	322	16	Q8YV66	Q8YV66 anabaena sp
567	6	2.8	295	12	Q9DWH0	Q9DWH0 rat cytoomeg	640	6	2.8	324	13	Q93549	Q93549 carassius a
568	6	2.8	297	2	Q93PPI	Q93PPI lactobacilli	641	6	2.8	325	16	Q9BMN4	Q9BMN4 rhizobium 1
569	6	2.8	297	5	Q22171	Q22171 caenorhabdi	642	6	2.8	325	16	Q819P8	Q819P8 listeria mo
570	6	2.8	297	5	Q9VE10	Q9VE10 drosophila	643	6	2.8	326	13	Q42171	Q42171 brachydanio
571	6	2.8	297	9	Q8SC28	Q8SC28 pseudomonas	644	6	2.8	327	2	Q9X4N3	Q9X4N3 bacteroides
572	6	2.8	297	11	Q99LH8	Q99LH8 mus musculu	645	6	2.8	327	16	Q9PEY7	Q9PEY7 xylella fas
573	6	2.8	297	16	Q97D12	Q97D12 clostridium	646	6	2.8	327	16	Q92WF7	Q92WF7 rhizobium m
574	6	2.8	297	17	Q9HR93	Q9HR93 halobacteri	647	6	2.8	328	13	Q91R35	Q91R35 cyprinus ca
575	6	2.8	298	17	Q50108	Q50108 pyrococcus	648	6	2.8	328	16	Q9RX33	Q9RX33 deinococcus
576	6	2.8	299	12	Q9WR74	Q9WR74 macaca mulla	649	6	2.8	329	2	Q44925	Q44925 borrelia bu
577	6	2.8	299	16	Q9BLK1	Q9BLK1 rhizobium 1	650	6	2.8	329	10	Q944U5	Q944U5 arabiidopsi
578	6	2.8	300	5	Q95TV3	Q95TV3 drosophila	651	6	2.8	329	16	Q8UJH8	Q8UJH8 agrobacteri
579	6	2.8	300	10	Q23226	Q23226 arabiidopsi	652	6	2.8	330	5	Q9Y023	Q9Y023 trypanosoma
580	6	2.8	300	13	Q8UWE7	Q8UWE7 tetradodon n	653	6	2.8	330	17	Q8ZTK0	Q8ZTK0 pyrobaculum
581	6	2.8	300	13	Q8UWP6	Q8UWP6 tetradodon n	654	6	2.8	331	3	Q8TRG2	Q8TRG2 schizosacch
582	6	2.8	302	2	Q24832	Q24832 actinobact	655	6	2.8	332	16	Q97SP2	Q97SP2 streptococ
583	6	2.8	302	5	Q9V6E1	Q9V6E1 drosophila	656	6	2.8	332	16	Q8XWL7	Q8XWL7 ralteconia s
584	6	2.8	303	3	Q9P7C1	Q9P7C1 schizosacch	657	6	2.8	333	10	Q9SY05	Q9SY05 arabiidopsi
585	6	2.8	303	16	Q928U8	Q928U8 listeria in	658	6	2.8	333	16	Q98KZ5	Q98KZ5 rhizobium 1
586	6	2.8	303	16	Q8Y4U2	Q8Y4U2 listeria mo	659	6	2.8	333	16	Q8RG17	Q8RG17 fusobacteri
587	6	2.8	304	4	Q96BP4	Q96BP4 homo sapien	660	6	2.8	334	16	Q9WYF5	Q9WYF5 thermotoga
588	6	2.8	304	10	Q940I6	Q940I6 arabiidopsi	661	6	2.8	334	16	Q97M28	Q97M28 clostridium
589	6	2.8	304	10	Q93VU4	Q93VU4 arabiidopsi	662	6	2.8	334	16	Q8XGB6	Q8XGB6 salmonella
590	6	2.8	304	10	Q8S3B4	Q8S3B4 arabiidopsi	663	6	2.8	335	4	Q15055	Q15055 homo sapien
591	6	2.8	305	2	Q9RHA2	Q9RHA2 thermus aqu	664	6	2.8	335	16	Q9ZBF9	Q9ZBF9 pseudomonas
592	6	2.8	305	10	Q41416	Q41416 solanum tub	665	6	2.8	335	17	Q97FV7	Q97FV7 streptomyce
593	6	2.8	305	17	Q8THK3	Q8THK3 methanosarc	666	6	2.8	336	4	Q14936	Q14936 homo sapien
594	6	2.8	306	8	Q8W770	Q8W770 saccharomyc	667	6	2.8	336	5	Q965G3	Q965G3 caenorhabdi
595	6	2.8	306	11	Q91VU1	Q91VU1 mus musculu	668	6	2.8	336	16	Q8RBQ3	Q8RBQ3 thermoaer
596	6	2.8	307	2	Q9Z157	Q9Z157 pseudomonas	669	6	2.8	337	17	Q9UX12	Q9UX12 sulfolobus
597	6	2.8	307	2	Q52653	Q52653 pseudomonas	670	6	2.8	337	16	Q87927	Q87927 borrelia bu
598	6	2.8	307	5	Q51962	Q51962 pseudomonas	671	6	2.8	338	5	Q02318	Q02318 caenorhabdi
599	6	2.8	307	5	Q18225	Q18225 caenorhabdi	672	6	2.8	338	11	Q922Q1	Q922Q1 mus musculu
600	6	2.8	307	16	Q8XKL0	Q8XKL0 clostridium	673	6	2.8	338	13	Q8QPS6	Q8QPS6 triticus ca

382	2.8	233	5	Q9N4L9	Q9n419 caenorhabdi	455	2.8	263	7	Q31163	Q31163 mus musculu
383	2.8	233	10	Q9FR80	Q9fr80 oryza sativ	456	2.8	263	17	Q9UYQ2	Q9uyq2 pyrococcus
384	2.8	234	16	Q8Y8Z9	Q8y8z9 listeria mo	457	2.8	264	7	Q78196	Q78196 mus musculu
385	2.8	238	12	Q90146	Q90146 macaca mula	458	2.8	264	7	Q31162	Q31162 mus musculu
386	2.8	238	16	Q8U9H8	Q8u9h8 agrobacteri	459	2.8	264	7	Q9T0A6	Q9tga6 rattus norv
387	2.8	238	17	Q29436	Q29436 archaeoglob	460	2.8	264	7	Q9TQAS	Q9tqs rattus norv
388	2.8	239	2	Q93UV7	Q93uv7 terzabacter	461	2.8	264	7	Q31103	Q31103 mus musculu
389	2.8	239	10	Q9M550	Q9m550 populus tom	462	2.8	264	7	Q31104	Q31104 mus musculu
390	2.8	239	10	Q24144	Q24144 nicotiana t	463	2.8	264	7	Q31159	Q31159 mus musculu
391	2.8	240	2	Q07457	Q07457 rhodospendo	464	2.8	264	16	Q911C9	Q911c9 pseudomonas
392	2.8	240	4	Q9H908	Q9h908 homo sapien	465	2.8	264	16	Q97RK9	Q97rk9 streptococ
393	2.8	240	10	Q02546	Q02546 nicotiana t	466	2.8	264	17	Q8U0L8	Q8u0l8 pyrococcus
394	2.8	240	10	Q04899	Q04899 nicotiana t	467	2.8	265	17	Q59221	Q59221 pyrococcus
395	2.8	241	10	Q43161	Q43161 stellaria l	468	2.8	265	17	Q9V120	Q9v120 pyrococcus
396	2.8	242	10	Q24149	Q24149 nicotiana t	469	2.8	267	11	Q9R1S4	Q9r1s4 rattus norv
397	2.8	242	10	Q24150	Q24150 nicotiana t	470	2.8	267	12	Q9WHD0	Q9whd0 kaposi's sa
398	2.8	242	10	Q24151	Q24151 nicotiana t	471	2.8	267	16	Q8XNLO	Q8xnlo clostridium
399	2.8	243	5	Q9N922	Q9n922 trypanosoma	472	2.8	268	5	Q9V3X2	Q9v3x2 drosophila
400	2.8	244	16	Q98C64	Q98c64 rhizobium l	473	2.8	268	11	Q8V8S2	Q8v8s2 mus musculu
401	2.8	244	16	Q97EL1	Q97el1 clostridium	474	2.8	268	17	Q8ZXQ7	Q8zxq7 pyrobaculu
402	2.8	244	17	Q30150	Q30150 archaeoglob	475	2.8	269	16	Q9F8Z1	Q9f8z1 streptomyce
403	2.8	245	5	Q17439	Q17439 bolentia vi	476	2.8	270	4	Q13670	Q13670 homo sapien
404	2.8	245	8	Q8WF52	Q8wf52 venerupis p	477	2.8	270	5	Q9G165	Q9gy65 leishmania
405	2.8	245	10	Q41720	Q41720 zinnia eleg	478	2.8	271	4	Q9H712	Q9h712 homo sapien
406	2.8	245	16	Q8RDH9	Q8rdh9 thermoanaer	479	2.8	271	16	Q8YNE5	Q8yne5 anaabaena sp
407	2.8	246	2	Q9FDM2	Q9fdm2 streptococ	480	2.8	271	17	Q973Z6	Q973z6 sulfolobus
408	2.8	246	10	Q8L185	Q8l185 eucalyptus	481	2.8	271	17	Q8TMS9	Q8tms9 methanopyru
409	2.8	246	16	Q9UV55	Q9jv55 neisseria m	482	2.8	272	2	Q07681	Q07681 escherichia
410	2.8	246	16	Q9H261	Q9h261 pseudomonas	483	2.8	272	5	Q9V5X6	Q9v5x6 drosophila
411	2.8	247	7	Q31124	Q31124 mus musculu	484	2.8	272	16	P96882	P96882 mycobacteri
412	2.8	247	10	Q40313	Q40313 medicago sa	485	2.8	272	17	Q9YA68	Q9ya68 aeropyrum p
413	2.8	247	10	Q42945	Q42945 nicotiana t	486	2.8	273	16	Q8ZJ74	Q8zj74 yerania pe
414	2.8	247	10	Q9SWM8	Q9swb8 eucalyptus	487	2.8	276	12	Q9MNS8	Q9mns8 kaposi's sa
415	2.8	247	10	Q9ATK7	Q9ack7 populus alb	488	2.8	276	12	Q9MNS6	Q9mns6 kaposi's sa
416	2.8	247	10	Q65922	Q65922 populus trt	489	2.8	278	10	Q9M265	Q9m265 araldopsis
417	2.8	247	10	Q65862	Q65862 populus trt	490	2.8	280	17	Q59139	P59139 caenorhabdi
418	2.8	247	16	Q9JX14	Q9jx14 neisseria m	491	2.8	280	17	Q8U142	Q8u142 pyrococcus
419	2.8	247	16	Q9XDL6	Q9xdl6 neisseria m	492	2.8	281	12	Q9WMH8	Q9wmh8 kaposi's sa
420	2.8	247	17	Q8TWM3	Q8twm3 methanogarc	493	2.8	281	12	Q9WMH9	Q9wmh9 kaposi's sa
421	2.8	249	2	Q9LBS5	Q9lbg5 geobacillus	494	2.8	281	16	Q91746	Q91746 pseudomonas
422	2.8	250	2	Q8RPW4	Q8rpw4 bacillus th	495	2.8	281	17	Q91746	Q91746 sulfolobus
423	2.8	250	16	Q98BC3	Q98bc3 rhizobium l	496	2.8	282	17	Q974N7	Q974n7 mycobacter
424	2.8	251	17	Q97CT4	Q97ct4 thermoplas	497	2.8	282	2	Q55068	Q55068 synechocyst
425	2.8	252	2	Q8VLH7	Q8vlh7 chlamydia p	498	2.8	282	8	Q47566	Q47566 agrocycbe ae
426	2.8	252	4	Q60817	Q60817 homo sapien	499	2.8	282	12	Q995B0	Q995b0 kaposi's sa
427	2.8	253	2	Q87827	Q87827 staphylococ	500	2.8	282	16	P73531	P73531 synechocyst
428	2.8	253	5	Q8SSP9	Q8ssp9 dictyosteli	501	2.8	282	16	P73976	P73976 synechocyst
429	2.8	253	16	Q9YV99	Q9yvn9 staphylococ	502	2.8	282	16	Q55645	Q55645 synechocyst
430	2.8	253	16	Q8XJ79	Q8xj79 clostridium	503	2.8	282	16	Q57248	Q57248 synechocyst
431	2.8	254	5	Q20788	Q20788 caenorhabdi	504	2.8	282	16	P75027	P75027 synechocyst
432	2.8	254	10	Q65162	Q65162 mesembryant	505	2.8	282	16	Q9PEY8	Q9pey8 xylella fas
433	2.8	254	16	Q24971	Q24971 helicobacte	506	2.8	282	16	Q9PEY6	Q9pey6 xylella fas
434	2.8	256	5	Q15692	Q15692 dictyosteli	507	2.8	282	16	Q97M19	Q97m19 clostridium
435	2.8	256	5	Q25081	Q25081 hypoderma l	508	2.8	282	16	Q9RR64	Q9rr64 diinococcus
436	2.8	258	12	Q9WHB5	Q9whb5 kaposi's sa	509	2.8	283	16	Q9PEY3	Q9pey3 xylella fas
437	2.8	258	16	Q98ID4	Q98id4 rhizobium l	510	2.8	283	16	Q9PEY2	Q9pey2 xylella fas
438	2.8	259	2	P97087	P97087 clostridium	511	2.8	283	16	Q9R1Y5	Q9r1y5 streptomyce
439	2.8	259	10	Q49499	Q49499 arabidopsis	512	2.8	283	17	Q8U1R2	Q8u1r2 pyrococcus
440	2.8	259	16	Q8REJ2	Q8rej2 fusobacteri	513	2.8	284	11	Q8VDP8	Q8vdp8 mus musculu
441	2.8	260	10	Q9XJ19	Q9xj19 oryza sativ	514	2.8	284	12	Q91GM3	Q91gm3 kaposi's sa
442	2.8	260	16	Q9AA01	Q9aa01 caulobacter	515	2.8	284	12	Q91GM2	Q91gm2 kaposi's sa
443	2.8	260	16	Q8RC92	Q8rc92 thermoanaer	516	2.8	284	12	Q91GM1	Q91gm1 kaposi's sa
444	2.8	260	17	Q9V172	Q9v172 pyrococcus	517	2.8	284	12	Q91GM0	Q91gm0 kaposi's sa
445	2.8	260	17	Q9V0U5	Q9v0u5 pyrococcus	518	2.8	284	12	Q91GV9	Q91gv9 kaposi's sa
446	2.8	261	4	Q9H9P7	Q9h9p7 homo sapien	519	2.8	284	12	Q91GV8	Q91gv8 kaposi's sa
447	2.8	261	7	Q9TQAT	Q9tqat rattus norv	520	2.8	284	12	Q91GV7	Q91gv7 kaposi's sa
448	2.8	261	12	Q91IH8	Q91ih8 kaposi's sa	521	2.8	284	12	Q91GV6	Q91gv6 kaposi's sa
449	2.8	261	16	Q9RYAT	Q9ryat deinococcus	522	2.8	284	12	Q91GV5	Q91gv5 kaposi's sa
450	2.8	261	16	Q92UM3	Q92uw3 rhizobium m	523	2.8	284	16	Q8RC06	Q8rc06 thermoanaer
451	2.8	262	12	Q9WH85	Q9wh85 kaposi's sa	524	2.8	285	12	Q9WMB4	Q9wmb4 kaposi's sa
452	2.8	262	12	Q9WH86	Q9wh86 kaposi's sa	525	2.8	285	16	Q92NA7	Q92na7 listeria in
453	2.8	262	12	Q9WH87	Q9wh87 kaposi's sa	526	2.8	285	16	Q9RK77	Q9rk77 streptomyce
454	2.8	262	12	Q9WH90	Q9wh90 kaposi's sa	527	2.8	286	17	Q28548	Q28548 archaeoglob

236	6	2.8	153	12	Q8V265	Q8V265 kapoti's sa	309	6	2.8	198	15	Q8UPG1	Q8UPG1 human immun
237	6	2.8	153	12	Q8V264	Q8V264 kapoti's sa	310	6	2.8	198	15	Q8UPG0	Q8UPG0 human immun
238	6	2.8	153	12	Q8V261	Q8V261 kapoti's sa	311	6	2.8	198	15	Q8UPF9	Q8UPF9 human immun
239	6	2.8	153	12	Q8V259	Q8V259 kapoti's sa	312	6	2.8	198	15	Q8UPF7	Q8UPF7 human immun
240	6	2.8	153	12	Q8V257	Q8V257 kapoti's sa	313	6	2.8	198	15	Q8UPF6	Q8UPF6 human immun
241	6	2.8	153	12	Q8V254	Q8V254 kapoti's sa	314	6	2.8	198	15	Q8UP19	Q8UP19 human immun
242	6	2.8	153	12	Q8V253	Q8V253 kapoti's sa	315	6	2.8	198	15	Q8UP18	Q8UP18 human immun
243	6	2.8	153	12	Q8V251	Q8V251 kapoti's sa	316	6	2.8	198	15	Q8UP08	Q8UP08 human immun
244	6	2.8	153	12	Q8V250	Q8V250 kapoti's sa	317	6	2.8	198	16	Q8XWD4	Q8XWD4 ralestonia s
245	6	2.8	153	12	Q8V249	Q8V249 kapoti's sa	318	6	2.8	199	17	Q86Z97	Q86Z97 sulfolobus
246	6	2.8	153	16	Q8R111	Q8R111 fusobacteri	319	6	2.8	200	12	Q85078	Q85078 poa semilat
247	6	2.8	154	5	Q9NAZ6	Q9NAZ6 ochlerotatu	320	6	2.8	200	13	Q9W6Y7	Q9W6Y7 rana escul
248	6	2.8	154	5	Q8UPF4	Q8UPF4 human immun	321	6	2.8	200	15	Q8UPF5	Q8UPF5 human immun
249	6	2.8	154	16	Q91470	Q91470 pseudomonas	322	6	2.8	200	15	Q8UP17	Q8UP17 human immun
250	6	2.8	156	10	Q81458	Q81458 arabiidopsis	323	6	2.8	200	15	Q8UP07	Q8UP07 human immun
251	6	2.8	157	16	Q9AC14	Q9AC14 caulobacter	324	6	2.8	201	2	Q9JWU6	Q9JWU6 buchiera ap
252	6	2.8	161	2	Q9F768	Q9F768 bacteroides	325	6	2.8	202	2	P96508	P96508 borrelia af
253	6	2.8	161	10	Q8V2V9	Q8V2V9 arabiidopsis	326	6	2.8	203	2	P96511	P96511 borrelia af
254	6	2.8	164	10	Q48879	Q48879 glycine max	327	6	2.8	204	10	Q94CS5	Q94CS5 oryza sativ
255	6	2.8	164	16	P73099	P73099 synechocyst	328	6	2.8	204	17	Q59145	Q59145 pyrococcus
256	6	2.8	164	16	Q56696	Q56696 synechocyst	329	6	2.8	207	3	Q01400	Q01400 neurospora
257	6	2.8	168	10	Q9FNM6	Q9FNM6 glycine tom	330	6	2.8	207	4	Q96038	Q96038 homo sapien
258	6	2.8	169	10	Q9LWD1	Q9LWD1 oryza sativ	331	6	2.8	207	10	Q9LUX9	Q9LUX9 arabiidopsis
259	6	2.8	169	16	Q8RBD2	Q8RBD2 thermoaer	332	6	2.8	207	16	Q9JXN2	Q9JXN2 streptococ
260	6	2.8	170	16	Q8RY07	Q8RY07 deinococcus	333	6	2.8	209	2	Q44671	Q44671 borrelia af
261	6	2.8	171	5	Q8S226	Q8S226 dirosophila	334	6	2.8	209	16	Q8ZFB4	Q8ZFB4 yersinia pe
262	6	2.8	171	10	Q9LR81	Q9LR81 arabiidopsis	335	6	2.8	210	15	Q76528	Q76528 human immun
263	6	2.8	171	16	P72636	P72636 synechocyst	336	6	2.8	210	16	Q84347	Q84347 chlamydia t
264	6	2.8	171	17	Q9V240	Q9V240 pyrococcus	337	6	2.8	211	2	Q49576	Q49576 borrelia af
265	6	2.8	172	4	Q9NPA9	Q9NPA9 homo sapien	338	6	2.8	211	16	Q8RE31	Q8RE31 fusobacteri
266	6	2.8	174	8	Q9B216	Q9B216 chrysomya c	339	6	2.8	211	17	Q97YW1	Q97YW1 sulfilobus
267	6	2.8	175	12	Q9DSH0	Q9DSH0 human rotav	340	6	2.8	212	2	Q49578	Q49578 borrelia af
268	6	2.8	175	16	Q31640	Q31640 bacillus su	341	6	2.8	212	2	Q44727	Q44727 borrelia bu
269	6	2.8	177	16	Q9CMQ4	Q9CMQ4 pasteurella	342	6	2.8	212	16	Q8Z008	Q8Z008 salmonella
270	6	2.8	178	2	Q93S18	Q93S18 streptococ	343	6	2.8	212	16	Q8Z718	Q8Z718 salmonella
271	6	2.8	178	10	Q39660	Q39660 chlorella v	344	6	2.8	213	17	Q58697	Q58697 pyrococcus
272	6	2.8	178	16	Q99R21	Q99R21 streptococ	345	6	2.8	214	2	Q9XDX3	Q9XDX3 bacillus me
273	6	2.8	178	16	Q66472	Q66472 aquifex aeo	346	6	2.8	216	10	Q9ZPR5	Q9ZPR5 arabiidopsis
274	6	2.8	179	10	Q9ZPZ4	Q9ZPZ4 arabiidopsis	347	6	2.8	217	11	Q8VCK2	Q8VCK2 mus musculu
275	6	2.8	179	16	Q32274	Q32274 bacillus su	348	6	2.8	217	12	Q90P64	Q90P64 hepatitis c
276	6	2.8	180	12	Q8V6L5	Q8V6L5 halovirus h	349	6	2.8	217	12	Q9J3F5	Q9J3F5 hepatitis c
277	6	2.8	181	4	Q9Y388	Q9Y388 homo sapien	350	6	2.8	218	2	Q93R38	Q93R38 bacillus ce
278	6	2.8	182	5	Q9XZYO	Q9XZYO leishmania	351	6	2.8	219	2	Q53027	Q53027 enterococc
279	6	2.8	182	17	Q8TW72	Q8TW72 methanopyru	352	6	2.8	219	16	Q98HA7	Q98HA7 rhizobium 1
280	6	2.8	184	2	Q53227	Q53227 rhodobacter	353	6	2.8	222	11	Q9CXV8	Q9CXV8 mus musculu
281	6	2.8	184	2	Q93P94	Q93P94 microscilla	354	6	2.8	223	12	Q991R2	Q991R2 soybean cri
282	6	2.8	185	12	Q9QBR0	Q9QBR0 feline cali	355	6	2.8	223	16	Q9KDE1	Q9KDE1 bacillus ha
283	6	2.8	186	10	Q9T057	Q9T057 arabiidopsis	356	6	2.8	223	16	Q9CF62	Q9CF62 lactococcus
284	6	2.8	186	11	Q8R185	Q8R185 mus musculu	357	6	2.8	223	16	Q984H8	Q984H8 rhizobium 1
285	6	2.8	187	16	Q8R104	Q8R104 fusobacteri	358	6	2.8	224	11	Q9C4L5	Q9C4L5 sulfilobus
286	6	2.8	188	5	Q18851	Q18851 caenorhabdi	359	6	2.8	224	11	Q99W71	Q99W71 mus musculu
287	6	2.8	188	16	Q97GJ0	Q97GJ0 clostridium	360	6	2.8	224	13	Q80G96	Q80G96 pagrus maio
288	6	2.8	189	2	Q9EZ23	Q9EZ23 rhizobium f	361	6	2.8	225	5	P90525	P90525 dictyosteli
289	6	2.8	189	2	Q53122	Q53122 mycobacteri	362	6	2.8	225	10	Q96Z62	Q96Z62 arabiidopsis
290	6	2.8	189	17	Q9ZPQ7	Q9ZPQ7 arabiidopsis	363	6	2.8	226	16	Q9RRL9	Q9RRL9 deinococcus
291	6	2.8	189	10	Q8ZXJ9	Q8ZXJ9 pyrobaculum	364	6	2.8	227	4	Q9BWA5	Q9BWA5 homo sapien
292	6	2.8	190	2	Q8VPS9	Q8VPS9 escherichia	365	6	2.8	227	10	Q9LZG7	Q9LZG7 arabiidopsis
293	6	2.8	190	16	Q8XSA9	Q8XSA9 ralerstonia s	366	6	2.8	227	12	Q9W866	Q9W866 equine arte
294	6	2.8	191	2	P94245	P94245 borrelia bu	367	6	2.8	228	17	Q9HUT2	Q9HUT2 thermoplas
295	6	2.8	192	2	Q50345	Q50345 lactobacilli	368	6	2.8	229	2	Q9EUV1	Q9EUV1 pseudomonas
296	6	2.8	192	5	Q22052	Q22052 caenorhabdi	369	6	2.8	229	12	Q9E3W4	Q9E3W4 tomato yell
297	6	2.8	192	5	Q02245	Q02245 caenorhabdi	370	6	2.8	229	12	Q96587	Q96587 avian adeno
298	6	2.8	193	10	Q04540	Q04540 arabiidopsis	371	6	2.8	229	16	Q8RE61	Q8RE61 fusobacteri
299	6	2.8	195	8	Q9T6K6	Q9T6K6 kluveromyc	372	6	2.8	230	12	Q902W4	Q902W4 barbados to
300	6	2.8	195	8	Q9T6K9	Q9T6K9 kluveromyc	373	6	2.8	230	16	Q9PKL3	Q9PKL3 chlamydia m
301	6	2.8	195	16	Q9L265	Q9L265 streptomyc	374	6	2.8	230	16	Q9A513	Q9A513 caulobacter
302	6	2.8	197	16	Q8RBJ0	Q8RBJ0 thermoaer	375	6	2.8	231	16	Q9PGH8	Q9PGH8 xylella fas
303	6	2.8	198	2	Q8VPV0	Q8VPV0 synechococ	376	6	2.8	231	16	Q98G59	Q98G59 rhizobium 1
304	6	2.8	198	4	Q96MM6	Q96MM6 homo sapien	377	6	2.8	231	16	Q9ZD04	Q9ZD04 listeria in
305	6	2.8	198	15	Q8UPG7	Q8UPG7 human immun	378	6	2.8	231	16	Q8V896	Q8V896 listeria mo
306	6	2.8	198	15	Q8UPG6	Q8UPG6 human immun	379	6	2.8	231	16	Q9LZP6	Q9LZP6 streptomyc
307	6	2.8	198	15	Q8UPG3	Q8UPG3 human immun	380	6	2.8	232	10	Q9SLP8	Q9SLP8 citrus nats
308	6	2.8	198	15	Q8UPG2	Q8UPG2 human immun	381	6	2.8	232	12	Q8QPR7	Q8QPR7 tomato molt



90	7	3.2	624	10	Q9M9M2	Q9M9M2 arabidopsis	163	6	2.8	109	5	Q26771	Q26771 trypanosoma
91	7	3.2	634	16	Q97H03	Q97H03 clostridium	164	6	2.8	109	16	Q98MA2	Q98MA2 rhizobium 1
92	7	3.2	660	10	Q49436	Q49436 arabidopsis	165	6	2.8	111	16	Q53492	Q53492 mycobacteri
93	7	3.2	662	10	Q9F061	Q9F061 brassica ol	166	6	2.8	112	2	O50328	O50328 escherichia
94	7	3.2	691	5	Q95V41	Q95V41 plasmodium	167	6	2.8	112	12	Q8V2G7	Q8V2G7 onion yello
95	7	3.2	699	5	Q77318	Q77318 plasmodium	168	6	2.8	113	5	Q9N923	Q9N923 trypanosoma
96	7	3.2	739	12	Q91BM8	Q91BM8 trichoplusi	169	6	2.8	115	16	Q8ZKX5	Q8ZKX5 salmonella
97	7	3.2	779	10	Q9CAP9	Q9CAP9 arabidopsis	170	6	2.8	115	17	Q97CU3	Q97CU3 thermoplas
98	7	3.2	803	16	Q8RAY9	Q8RAY9 thermoaer	171	6	2.8	118	16	Q9PQ20	Q9PQ20 ureaplasma
99	7	3.2	816	11	Q9CS25	Q9CS25 mus musculu	172	6	2.8	120	7	Q31186	Q31186 mus musculu
100	7	3.2	837	5	Q25751	Q25751 plasmodium	173	6	2.8	120	7	Q31105	Q31105 mus musculu
101	7	3.2	852	4	Q9H956	Q9H956 homo sapien	174	6	2.8	121	4	Q9H246	Q9H246 homo sapien
102	7	3.2	866	16	Q8Y570	Q8Y570 listeria mo	175	6	2.8	121	16	Q99W88	Q99W88 staphylococ
103	7	3.2	985	5	Q9NHD2	Q9NHD2 maecina domes	176	6	2.8	121	16	Q8R144	Q8R144 fusobacteri
104	7	3.2	986	5	Q9N9M3	Q9N9M3 haemonchus	177	6	2.8	123	16	Q9HUX0	Q9HUX0 pseudomonas
105	7	3.2	1011	10	Q9XEA4	Q9XEA4 arabidopsis	178	6	2.8	123	16	Q97HS8	Q97HS8 clostridium
106	7	3.2	1059	11	Q922B7	Q922B7 mus musculu	179	6	2.8	124	17	Q9HVL9	Q9HVL9 thermoplas
107	7	3.2	1077	10	Q81441	Q81441 arabidopsis	180	6	2.8	126	16	Q8R6J6	Q8R6J6 fusobacteri
108	7	3.2	1122	10	Q93W78	Q93W78 oryza sativ	181	6	2.8	129	16	Q8ZID2	Q8ZID2 yersinia pe
109	7	3.2	1143	11	Q923A3	Q923A3 mus musculu	182	6	2.8	130	10	Q9AYD5	Q9AYD5 oryza sativ
110	7	3.2	1206	16	Q9X1T0	Q9X1T0 thermotoga	183	6	2.8	130	16	Q8UDW0	Q8UDW0 agrobacteri
111	7	3.2	1392	11	Q8ROF2	Q8ROF2 mus musculu	184	6	2.8	131	4	Q9H098	Q9H098 homo sapien
112	7	3.2	1401	4	Q15021	Q15021 homo sapien	185	6	2.8	131	8	Q8W908	Q8W908 mesocriceta
113	7	3.2	1426	3	P78577	P78577 emeticella	186	6	2.8	131	11	Q9CR82	Q9CR82 mus musculu
114	7	3.2	1472	13	Q90ZAO	Q90ZAO gallus gall	187	6	2.8	131	16	Q97G60	Q97G60 clostridium
115	7	3.2	1490	5	Q9TX96	Q9TX96 drosophila	188	6	2.8	132	16	Q9PIC0	Q9PIC0 campylobact
116	7	3.2	1550	16	Q8YX08	Q8YX08 anabaena sp	189	6	2.8	133	2	Q91UT1	Q91UT1 rhizobium m
117	7	3.2	1603	5	Q21884	Q21884 caenorhabdi	190	6	2.8	135	7	Q9BC18	Q9BC18 rattus norv
118	7	3.2	1821	16	Q8YEP6	Q8YEP6 anabaena sp	191	6	2.8	135	16	Q8XD42	Q8XD42 escherichia
119	7	3.2	1822	2	Q07290	Q07290 streptococc	192	6	2.8	136	12	Q91IH2	Q91IH2 kaposi's sa
120	7	3.2	2362	5	Q9VLA0	Q9VLA0 drosophila	193	6	2.8	136	16	Q8RHZ6	Q8RHZ6 fusobacteri
121	7	3.2	4345	5	Q9U459	Q9U459 plasmodium	194	6	2.8	137	2	Q93MM6	Q93MM6 lactobacilli
122	7	3.2	5458	5	Q9NFS3	Q9NFS3 drosophila	195	6	2.8	137	5	Q24960	Q24960 giardia lam
123	7	3.2	16215	5	Q9BM56	Q9BM56 lineus sp.	196	6	2.8	137	16	Q8ZM81	Q8ZM81 salmonella
124	6	2.8	25	7	Q19447	Q19447 mus musculu	197	6	2.8	139	2	Q93UD9	Q93UD9 carsonella
125	6	2.8	30	7	Q19447	Q19447 mus musculu	198	6	2.8	139	2	Q93UD4	Q93UD4 carsonella
126	6	2.8	30	7	Q19448	Q19448 mus musculu	199	6	2.8	139	2	Q93UB8	Q93UB8 carsonella
127	6	2.8	39	16	Q9A7N6	Q9A7N6 caulobacter	200	6	2.8	139	2	Q93UB2	Q93UB2 carsonella
128	6	2.8	46	16	Q8ZMY4	Q8ZMY4 salmonella	201	6	2.8	139	2	Q93UA3	Q93UA3 carsonella
129	6	2.8	50	11	Q924X5	Q924X5 mus musculu	202	6	2.8	139	2	Q93U97	Q93U97 carsonella
130	6	2.8	64	9	Q9AYX2	Q9AYX2 lactococcus	203	6	2.8	139	2	Q93UB8	Q93UB8 carsonella
131	6	2.8	66	10	Q8W542	Q8W542 retama raet	204	6	2.8	139	2	Q93UB2	Q93UB2 carsonella
132	6	2.8	69	5	Q95XP7	Q95XP7 caenorhabdi	205	6	2.8	139	2	Q93U67	Q93U67 carsonella
133	6	2.8	75	16	Q8R8P1	Q8R8P1 thermoaer	206	6	2.8	139	2	Q93U67	Q93U67 carsonella
134	6	2.8	77	3	Q74989	Q74989 schizosacch	207	6	2.8	139	2	Q93U64	Q93U64 carsonella
135	6	2.8	78	5	Q9NAU1	Q9NAU1 mytilus edu	208	6	2.8	139	2	Q93U18	Q93U18 carsonella
136	6	2.8	79	13	Q98TF3	Q98TF3 platyichthys	209	6	2.8	139	17	Q59018	Q59018 methanococc
137	6	2.8	79	16	Q9A188	Q9A188 streptococc	210	6	2.8	140	4	Q9AM41	Q9AM41 yersinia ps
138	6	2.8	80	9	Q38489	Q38489 bacterioph	211	6	2.8	140	16	Q96LP1	Q96LP1 homo sapien
139	6	2.8	86	9	Q8W5Y9	Q8W5Y9 bacterioph	212	6	2.8	140	16	Q8ZFC2	Q8ZFC2 yersinia pe
140	6	2.8	86	16	Q92B48	Q92B48 listeria in	213	6	2.8	141	5	Q26709	Q26709 trypanosoma
141	6	2.8	87	11	Q61713	Q61713 mus musculu	214	6	2.8	141	16	Q9KCU3	Q9KCU3 bacillus ha
142	6	2.8	88	10	Q9ZRN1	Q9ZRN1 melilotus a	215	6	2.8	142	12	Q91IH5	Q91IH5 kaposi's sa
143	6	2.8	88	13	Q9OWP9	Q9OWP9 pleuroctes a	216	6	2.8	142	12	Q91IH5	Q91IH5 kaposi's sa
144	6	2.8	88	16	Q98P21	Q98P21 mycoplasma	217	6	2.8	144	3	Q96WQ5	Q96WQ5 pleurotus c
145	6	2.8	89	10	Q9ZPR9	Q9ZPR9 arabidopsis	218	6	2.8	144	12	Q91IH3	Q91IH3 kaposi's sa
146	6	2.8	89	10	Q9ZRN0	Q9ZRN0 melilotus o	219	6	2.8	145	16	Q8ZRR3	Q8ZRR3 salmonella
147	6	2.8	90	10	Q9ZRN2	Q9ZRN2 genista tin	220	6	2.8	145	16	Q8ZRD5	Q8ZRD5 salmonella
148	6	2.8	91	2	Q56426	Q56426 thermus the	221	6	2.8	146	16	Q9AB17	Q9AB17 caulobacter
149	6	2.8	92	12	Q90397	Q90397 hepatitis g	222	6	2.8	147	16	Q8RCV5	Q8RCV5 thermoaer
150	6	2.8	93	16	Q92DU5	Q92DU5 listeria in	223	6	2.8	147	17	Q97B48	Q97B48 thermoplas
151	6	2.8	93	16	Q8Y932	Q8Y932 listeria mo	224	6	2.8	148	5	Q9GQX0	Q9GQX0 paramphisto
152	6	2.8	94	10	Q93M18	Q93M18 cornebacte	225	6	2.8	148	12	Q91IG9	Q91IG9 kaposi's sa
153	6	2.8	94	10	Q9LXG5	Q9LXG5 arabidopsis	226	6	2.8	148	16	Q53397	Q53397 mycobacteri
154	6	2.8	97	16	P97152	P97152 eynchoyest	227	6	2.8	149	2	Q93BD5	Q93BD5 mycobacteri
155	6	2.8	101	16	Q9PH51	Q9PH51 xyella fas	228	6	2.8	150	2	Q86275	Q86275 lactococcus
156	6	2.8	102	2	Q45390	Q45390 bordetella	229	6	2.8	150	5	Q9VNZ0	Q9VNZ0 drosophila
157	6	2.8	102	16	Q8R6M4	Q8R6M4 thermoaer	230	6	2.8	152	16	Q8YX00	Q8YX00 anabaena sp
158	6	2.8	104	17	Q9HRD2	Q9HRD2 halobacteri	231	6	2.8	152	16	Q8XOA9	Q8XOA9 ralsstromia s
159	6	2.8	107	10	Q9FWU0	Q9FWU0 tulipa gesm	232	6	2.8	153	12	Q8V269	Q8V269 kaposi's sa
160	6	2.8	107	16	Q9K9Q2	Q9K9Q2 bacillus ha	233	6	2.8	153	12	Q8V268	Q8V268 kaposi's sa
161	6	2.8	109	2	Q939U2	Q939U2 rhodovulum	234	6	2.8	153	12	Q8V267	Q8V267 kaposi's sa
162	6	2.8	109	4	Q9WQ01	Q9WQ01 homo sapien	235	6	2.8	153	12	Q8V266	Q8V266 kaposi's sa



GenCore version 5.1.4 pj\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 9, 2003, 14:39:23 ; Search time 85 Seconds  
(without alignments)  
528.450 Million cell updates/sec

Title: us-10-034-500-2

Perfect score: 218 1 AEVTASCTKVESYNYLVY.....QEQLVTEVVVLRGNFAP 218

Sequence:

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size : 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

1: SPTRMBL\_21:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mammal:\*  
8: sp\_mmc:\*  
9: sp\_organelle:\*  
10: sp\_phase:\*  
11: sp\_plant:\*  
12: sp\_rodent:\*  
13: sp\_virus:\*  
14: sp\_vertebrate:\*  
15: sp\_unclassified:\*  
16: sp\_virus:\*  
17: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	3.7	210	2	Q45178
2	7	3.2	120	2	Q47616
3	7	3.2	120	10	Q8RVW9
4	7	3.2	120	10	Q8RVW8
5	7	3.2	120	10	Q8RVW7
6	7	3.2	140	10	Q9SHZ7
7	7	3.2	157	4	015401
8	7	3.2	159	4	Q9R7B0
9	7	3.2	163	2	Q9R7A9
10	7	3.2	175	2	Q9R7B4
11	7	3.2	176	2	Q34119
12	7	3.2	177	2	Q44999
13	7	3.2	178	2	Q9REH7
14	7	3.2	180	2	Q34123
15	7	3.2	182	2	Q9R7B2
16	7	3.2	191	2	P94223

17	7	3.2	191	2	P70818	P70818 borrelia bu
18	7	3.2	192	2	P94243	P94243 borrelia bu
19	7	3.2	192	2	Q9S3P3	Q9S3P3 borrelia bu
20	7	3.2	192	2	Q9R7B1	Q9R7B1 borrelia bu
21	7	3.2	193	2	Q9R7B3	Q9R7B3 borrelia bu
22	7	3.2	193	2	Q31115	Q31115 borrelia bu
23	7	3.2	193	2	Q9S317	Q9S317 borrelia ga
24	7	3.2	194	2	Q9S516	Q9S516 borrelia ga
25	7	3.2	196	2	Q44997	Q44997 borrelia ga
26	7	3.2	197	16	Q9PC51	Q9PC51 xylella fas
27	7	3.2	200	2	Q9R0R9	Q9R0R9 borrelia bu
28	7	3.2	200	2	Q93098	Q93098 borrelia bu
29	7	3.2	202	2	Q932W2	Q932W2 borrelia bu
30	7	3.2	202	16	Q8Y1B5	Q8Y1B5 bruceia me
31	7	3.2	203	2	Q44983	Q44983 borrelia ga
32	7	3.2	204	2	Q44984	Q44984 borrelia ga
33	7	3.2	205	2	Q50622	Q50622 borrelia af
34	7	3.2	205	2	Q08141	Q08141 borrelia bu
35	7	3.2	209	2	Q45179	Q45179 borrelia bu
36	7	3.2	209	2	Q9K1K3	Q9K1K3 borrelia ga
37	7	3.2	210	2	Q9K1M6	Q9K1M6 borrelia ga
38	7	3.2	210	2	Q45647	Q45647 borrelia ga
39	7	3.2	210	2	Q49582	Q49582 borrelia ga
40	7	3.2	210	2	Q57279	Q57279 borrelia ga
41	7	3.2	212	2	Q49580	Q49580 borrelia ga
42	7	3.2	214	2	Q44646	Q44646 borrelia af
43	7	3.2	220	11	Q9DB47	Q9DB47 mus musculus
44	7	3.2	223	16	Q92V41	Q92V41 rhizobium m
45	7	3.2	237	4	Q9UHL1	Q9UHL1 homo sapien
46	7	3.2	246	4	Q75586	Q75586 homo sapien
47	7	3.2	246	4	Q9BTH1	Q9BTH1 homo sapien
48	7	3.2	251	16	Q8RDD7	Q8RDD7 thermotoma
49	7	3.2	252	16	Q9S211	Q9S211 streptomyces
50	7	3.2	257	17	Q9UZY6	Q9UZY6 pyrococcus
51	7	3.2	266	16	Q9KCP7	Q9KCP7 pseudomonas
52	7	3.2	267	16	Q9KCP4	Q9KCP4 bacillus ha
53	7	3.2	276	17	Q9HPN0	Q9HPN0 halobacter
54	7	3.2	304	16	Q92286	Q92286 rhizobium m
55	7	3.2	306	10	Q9FL74	Q9FL74 arbidopsis
56	7	3.2	309	2	Q9ZFP2	Q9ZFP2 enterococcus
57	7	3.2	311	16	Q910U8	Q910U8 pseudomonas
58	7	3.2	314	11	Q8VEM8	Q8VEM8 mus musculus
59	7	3.2	317	10	Q93WU5	Q93WU5 gossypium h
60	7	3.2	338	8	Q9TLE2	Q9TLE2 cyanidium c
61	7	3.2	338	16	Q99P21	Q99P21 streptomyces
62	7	3.2	343	16	Q8X7Z9	Q8X7Z9 escherichia
63	7	3.2	351	16	Q8RTB8	Q8RTB8 fusobacteri
64	7	3.2	355	16	Q9A825	Q9A825 caulobacter
65	7	3.2	357	2	Q9X9Q0	Q9X9Q0 streptomyces
66	7	3.2	357	2	Q8RNX2	Q8RNX2 streptomyces
67	7	3.2	358	10	Q9S7S1	Q9S7S1 arbidopsis
68	7	3.2	367	16	Q9ADG7	Q9ADG7 streptomyces
69	7	3.2	379	10	Q49447	Q49447 arbidopsis
70	7	3.2	394	16	Q8R965	Q8R965 thermotoma
71	7	3.2	402	10	Q8RWM3	Q8RWM3 arbidopsis
72	7	3.2	402	16	Q8RHF1	Q8RHF1 fusobacteri
73	7	3.2	416	2	Q31191	Q31191 myxococcus
74	7	3.2	422	9	Q9X1J2	Q9X1J2 bacterioph
75	7	3.2	422	16	Q9KXB5	Q9KXB5 escherichia
76	7	3.2	429	16	Q25470	Q25470 helicobacte
77	7	3.2	429	16	Q9ZL62	Q9ZL62 blumeria gr
78	7	3.2	430	3	Q96VW2	Q96VW2 mytilus edu
79	7	3.2	441	5	Q8T6U5	Q8T6U5 mytilus edu
80	7	3.2	444	5	Q8T5C3	Q8T5C3 mytilus edu
81	7	3.2	447	16	Q8Y1F0	Q8Y1F0 anabaena sp
82	7	3.2	453	5	Q8T5C2	Q8T5C2 mytilus gal
83	7	3.2	487	2	Q52719	Q52719 klebsiella
84	7	3.2	518	5	Q76345	Q76345 homarus ame
85	7	3.2	524	10	Q9Z0J3	Q9Z0J3 plasmodium
86	7	3.2	545	5	Q44019	Q44019 arbidopsis
87	7	3.2	551	10	Q9S1I8	Q9S1I8 arbidopsis
88	7	3.2	551	10	Q94C51	Q94C51 arbidopsis
89	7	3.2	606	10	Q9FT74	Q9FT74 arbidopsis

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 9, 2003, 14:33:43 ; Search time 25 Seconds  
(without alignments)  
361.674 Million cell updates/sec

Title: US-10-034-500-2  
Perfect score: 218  
Sequence: 1 AEVTASCTKRVESYNLYVDY.....QEQILVTEEVVLGRVNFAR 218

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size : 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	3.2	210	1	OSCL_BOBBU
2	7	3.2	213	1	HIS1_PSEAE
3	7	3.2	343	1	ABC_ECOLI
4	7	3.2	390	1	Y4TL_RHISN
5	7	3.2	394	1	PGK_THENY
6	7	3.2	401	1	PGK_SYNY3
7	7	3.2	482	1	SYP_METHN
8	7	3.2	552	1	VNS1_BTVIA
9	7	3.2	552	1	VNS1_BTV20
10	7	3.2	553	1	G6PI_YEAST
11	7	3.2	561	1	YBUL_ECOLI
12	7	3.2	630	1	Y4BJ_RHISN
13	6	2.8	15	1	ESTU_MANSE
14	6	2.8	68	1	UCRH_SOLTU
15	6	2.8	93	1	REPOL_SUPTO
16	6	2.8	100	1	YE26_AQUAE
17	6	2.8	106	1	GLRX_RABIT
18	6	2.8	121	1	FOLB_STNAU
19	6	2.8	124	1	PRO2_DICDI
20	6	2.8	132	1	NUSB_CAMJE
21	6	2.8	135	1	RNH1_BACSU
22	6	2.8	135	1	YGFY_ECOLI
23	6	2.8	140	1	RK16_CVACA
24	6	2.8	141	1	YE17_YEAST
25	6	2.8	141	1	YHBC_HABIN
26	6	2.8	146	1	HBB2_XENLA
27	6	2.8	147	1	GLB_PAREP
28	6	2.8	150	1	FLAG_MERYO
29	6	2.8	150	1	RL9_STRBN
30	6	2.8	152	1	YKH2_CAREL
31	6	2.8	153	1	LG81_LUPLU
32	6	2.8	153	1	LG82_LUPLU
33	6	2.8	153	1	NUSB_FUSNN
34	6	2.8	154	1	YBEY_HAEIN
35	6	2.8	156	1	RSS_BUCKA
36	6	2.8	175	1	IL10_PIG
37	6	2.8	184	1	REGA_RHOSH
38	6	2.8	202	1	FGF4_MOUSE
39	6	2.8	203	1	INFB_CHICK
40	6	2.8	206	1	BCHU_RHOSH
41	6	2.8	210	1	END3_BUCAI
42	6	2.8	213	1	YCFM_ECOLI
43	6	2.8	216	1	MERB_STNAU
44	6	2.8	218	1	MERB_BACSR
45	6	2.8	225	1	Y409_MYCPN
46	6	2.8	226	1	BASP_HUMAN
47	6	2.8	233	1	LEC_VICPA
48	6	2.8	239	1	YQAS_BACSU
49	6	2.8	242	1	CAMT_VITVI
50	6	2.8	245	1	FRDB_HELPJ
51	6	2.8	245	1	FRDB_HELPJ
52	6	2.8	247	1	CAMT_POPTM
53	6	2.8	249	1	LEC2_CYTSC
54	6	2.8	249	1	CAMT_EUCGU
55	6	2.8	249	1	COX2_SACEX
56	6	2.8	249	1	GPP1_YEAST
57	6	2.8	249	1	RS2_HALNT
58	6	2.8	250	1	YZ52_MYCTU
59	6	2.8	253	1	TPIS_BACSU
60	6	2.8	256	1	HYPB_HYPLI
61	6	2.8	258	1	YDP2_SCHPO
62	6	2.8	264	1	HB21_MOUSE
63	6	2.8	264	1	HB22_MOUSE
64	6	2.8	264	1	HB2D_MOUSE
65	6	2.8	264	1	HB2D_RAT
66	6	2.8	264	1	HB2J_MOUSE
67	6	2.8	264	1	HB2J_MOUSE
68	6	2.8	268	1	CPC2_SYNP2
69	6	2.8	274	1	OMP2_CHLPN
70	6	2.8	279	1	PANE_PYRAE
71	6	2.8	282	1	ARO2_MERYA
72	6	2.8	282	1	PANE_PYRAB
73	6	2.8	283	1	PANE_PYRHO
74	6	2.8	286	1	YOL8_PSEAE
75	6	2.8	290	1	ARV1_MESAU
76	6	2.8	290	1	ARV1_RAT
77	6	2.8	290	1	ARV2_RAT
78	6	2.8	290	1	LEG2_CIALU
79	6	2.8	293	1	YC10_ECOLI
80	6	2.8	294	1	SR22_SAREP
81	6	2.8	298	1	FLUG_BACSU
82	6	2.8	298	1	YAKI_YEAST
83	6	2.8	302	1	YCAN_ECOLI
84	6	2.8	304	1	P32_MYCGA
85	6	2.8	307	1	Y123_MYCTU
86	6	2.8	310	1	DP3B_MICLU
87	6	2.8	310	1	URB1_RAT
88	6	2.8	312	1	ACCA_HELPJ
89	6	2.8	312	1	ACCA_HELPJ
90	6	2.8	314	1	TPIC_PRAAN
91	6	2.8	317	1	L1P1_PSYIM
92	6	2.8	317	1	MTB1_NETGO
93	6	2.8	318	1	MTB2_HAEAE
94	6	2.8	319	1	OTCA_BACSU
95	6	2.8	319	1	THIL_MERYA
96	6	2.8	320	1	HEP2_BACST
97	6	2.8	320	1	SIAL_RAT
98	6	2.8	321	1	YDG7_SCHPO
99	6	2.8	322	1	MT04_DEIRA
100	6	2.8	325	1	YJ40_YEAST
101	6	2.8	329	1	ESG2_TRYBB
102	6	2.8	330	1	Y367_MERYA
103	6	2.8	333	1	KC23_ARATH
104	6	2.8	334	1	GPDA_CHUTR
105	6	2.8	334	1	HFLC_ECOLI
106	6	2.8	338	1	BZTA_RHOCA
P71335	haemophilus				
P46183	buchnera ap				
Q29055	sus scrofa				
Q53228	rhodobacter				
P11403	mus musculus				
Q90873	gallus gall				
Q92567	rhodobacter				
P57219	buchnera ap				
P75947	escherichia				
P08653	staphylococ				
P16172	bacillus sp				
P75187	mycoplasma				
P80723	homo sapien				
P02871	vicia faba				
P45915	bacillus su				
Q43237	vicia vintif				
Q92mp1	helicobacte				
O06914	helicobacte				
Q43095	populus tre				
P29257	cytilus sco				
O04854	eucalyptus				
P43377	saccharomyc				
P41277	saccharomyc				
P57713	halobacteri				
P71848	mycobacteri				
P27876	bacillus su				
P35568	hypodermia 1				
O10474	escheriasch				
P04230	mus musculus				
P01915	mus musculus				
P20040	mus musculus				
P18211	rattus norv				
P18468	mus musculus				
P18469	mus musculus				
P31967	synechococ				
Q92616	chlamydia p				
Q82170	pyrobaculum				
O58484	methanococ				
Q9V061	pyrococcus				
O58665	pyrococcus				
Q91153	pseudomonas				
P50292	mesocricetu				
P50297	rattus norv				
Q39599	cladrasia				
P72653	escherichia				
P24489	sarcophaga				
P96501	bacillus su				
P39551	saccharomyc				
P75816	escherichia				
O49378	mycoplasma				
O50610	mycobacteri				
P21174	micrococcus				
P51533	rattus norv				
Q92523	helicobacte				
O94489	fragaria an				
O02104	psychrobact				
Q59603	neisseria g				
O30868	haemophilus				
P18186	bacillus su				
O60319	methanococ				
P55757	methanococ				
P13899	rattus norv				
Q10474	echinosach				
Q91666	deinococcus				
P47120	saccharomyc				
P04478	crypanosoma				
O57813	methanococ				
O64817	arabidopsis				
O84719	chlamydia t				
P25661	escherichia				
O52663	rhodobacter				

107	6	2.8	344	1	LIC1_BURCE	P22089	burkholderi	180	6	2.8	465	1	SMAS_MOUSE	P97454	mus musculus
108	6	2.8	344	1	LICH_PSEB5	P25276	pseudomonas	181	6	2.8	465	1	SMAS_RAT	Q9173	rattus norv
109	6	2.8	344	1	SERB_ARCFU	O28142	archaeoglob	182	6	2.8	468	1	SMAL_RAT	P97568	rattus norv
110	6	2.8	345	1	TRPD_SUTSO	P50384	suifolobus	183	6	2.8	472	1	SMAL_BRARE	Q91872	brachydanio
111	6	2.8	348	1	PLSX_SYNA3	P73950	synecocyst	184	6	2.8	472	1	SPSB_BACSU	P39622	baecillus su
112	6	2.8	349	1	ISPG_CLOAB	O9156	clostridium	185	6	2.8	472	1	XYIA_ARATH	O9167	arabidopsis
113	6	2.8	351	1	LEPD_RHIL0	O98mc4	rhizobium	186	6	2.8	475	1	CYSN_ECOLI	P23845	eaccherichia
114	6	2.8	356	1	DDL_STYAM	O9984	staphylococ	187	6	2.8	475	1	2131_HUMAN	P27739	homo sapien
115	6	2.8	357	1	LEU5_YEAST	P38702	saccharomyc	188	6	2.8	486	1	CLS_ECOLI	P31071	eaccherichia
116	6	2.8	357	1	VAL1_TYLCV	P27259	tomato yell	189	6	2.8	486	1	MURE_BACHD	Q9K964	baecillus ha
117	6	2.8	358	1	VAL1_CLVX	P14982	caasava lat	190	6	2.8	489	1	ACCD_BRANA	P48937	brasica na
118	6	2.8	358	1	VAL1_CLVX	P14972	caasava lat	191	6	2.8	489	1	P40210	saccharomyc	
119	6	2.8	361	1	LEU1_PYRHO	O59390	pyrococcus	192	6	2.8	490	1	TMS2_MOUSE	Q9116	mus musculus
120	6	2.8	363	1	QUEA_PASNU	P57832	pasteurella	193	6	2.8	492	1	TMS2_HUMAN	O15353	homo sapien
121	6	2.8	365	1	ISPG_PROST	P72241	providencia	194	6	2.8	504	1	DEGP_RHIME	O59337	deinococcus
122	6	2.8	366	1	ALR_STRPY	O99Y98	strepococc	195	6	2.8	511	1	LGT_CITUN	Q9mb73	citrus unsh
123	6	2.8	366	1	APRI_SCHRO	P50525	schizosacch	196	6	2.8	512	1	PNTA_HATIN	P43842	haemophilus
124	6	2.8	367	1	ALR_STRPN	O54899	strepococc	197	6	2.8	513	1	AMY3_SCHPO	O14154	schizosacch
125	6	2.8	370	1	ISPG_BACHD	O94d18	baecillus ha	198	6	2.8	523	1	MWSA_CAEEL	P22713	caenorhabdi
126	6	2.8	370	1	AROB_CAUOR	O94434	caulobacter	199	6	2.8	533	1	MWIN_RHITR	O05467	rhizobium t
127	6	2.8	371	1	OOB2_RHIME	P72299	rhizobium m	200	6	2.8	536	1	CATA_DEIRA	O59337	deinococcus
128	6	2.8	371	1	OOB2_AGRF4	O59159	agrobacteri	201	6	2.8	547	1	CATX_BACSU	P4377	baecillus su
129	6	2.8	372	1	DHA_BACST	P17557	baecillus st	202	6	2.8	548	1	YDD2_SCHPO	O10428	schizosacch
130	6	2.8	372	1	VSGF_EBOSB	O84455	ebola virus	203	6	2.8	552	1	VNS1_BTW10	P07131	bluetongue
131	6	2.8	375	1	ISPG_YERPE	P58672	yersinia pe	204	6	2.8	552	1	VNS1_BTW11	P55932	bluetongue
132	6	2.8	377	1	ISPG_BACSU	P54482	baecillus su	205	6	2.8	552	1	VNS1_BTW13	P35933	bluetongue
133	6	2.8	380	1	MNSA_XENLA	P31286	xenopus lae	206	6	2.8	552	1	VNS1_BTW17	P14245	bluetongue
134	6	2.8	384	1	YE48_ANASP	P29978	anabaena sp	207	6	2.8	552	1	VNS1_BTW1S	P33472	bluetongue
135	6	2.8	386	1	SCW4_YEAST	P53334	saccharomyc	208	6	2.8	552	1	VNS1_BTW2A	P35931	bluetongue
136	6	2.8	387	1	GAT6_CHICK	P4693	gallus gall	209	6	2.8	553	1	YMS8_YEAST	P55931	bluetongue
137	6	2.8	387	1	ISPG_MYCTU	O33350	mycobacteri	210	6	2.8	555	1	GHT3_SCHPO	O04847	saccharomyc
138	6	2.8	387	1	PKNS_MYCGE	P47355	mycoplasma	211	6	2.8	556	1	HUTU_PSEPU	P25080	pseudomonas
139	6	2.8	392	1	BM1S_HUMAN	O95972	homo sapien	212	6	2.8	557	1	SVR_BACHD	Q9K6c1	baecillus ha
140	6	2.8	392	1	CEB2_RANVA	O93229	rana japoni	213	6	2.8	557	1	G6PI_RABIT	O9K6c2	oryctolagus
141	6	2.8	392	1	CEB2_XENLA	P13351	xenopus lae	214	6	2.8	557	1	GHT4_SCHPO	O59932	schizosacch
142	6	2.8	394	1	PGK_BACME	P24269	baecillus me	215	6	2.8	559	1	HUTU_PSEAE	O9h83	pseudomonas
143	6	2.8	394	1	PGK_BACST	P18912	baecillus st	216	6	2.8	562	1	YMX8_YEAST	O04301	saccharomyc
144	6	2.8	394	1	FLGE_BRUBA	P40924	baecillus su	217	6	2.8	565	1	HUTU_PSEBX	Q10574	pseudomonas
145	6	2.8	396	1	FLGE_BRUBA	O52070	bruceella ab	218	6	2.8	570	1	SP26_CAEEL	Q10579	caenorhabdi
146	6	2.8	399	1	DP3B_MYCLE	P46887	mycobacteri	219	6	2.8	573	1	PTI_BORBU	O51508	borreliia bu
147	6	2.8	403	1	YD87_SCHPO	O10412	schizosacch	220	6	2.8	584	1	DNL1_PYRAE	O93377	pyrobaculum
148	6	2.8	404	1	HCGY_HELPO	P56823	helix pomat	221	6	2.8	591	1	UI49_EBV	P14347	epstein-bar
149	6	2.8	405	1	S3AE_BACSU	P49782	baecillus su	222	6	2.8	598	1	YUOF_SCHPO	Q10168	schizosacch
150	6	2.8	406	1	SNX6_HUMAN	Q9uh77	homo sapien	223	6	2.8	600	1	SP08_YEAST	P41833	saccharomyc
151	6	2.8	407	1	PEPT_STRPY	Q9a6f4	strepococc	224	6	2.8	615	1	NIFA_AZOCA	P09133	azorhizobiu
152	6	2.8	407	1	YK67_MYCTU	O10678	mycobacteri	225	6	2.8	622	1	COX1_BACSU	P44010	baecillus su
153	6	2.8	408	1	PGK_SUTSO	P50317	sulfolobus	226	6	2.8	628	1	VZ68_PSEAE	P28812	pseudomonas
154	6	2.8	410	1	CATD_MOUSE	P18242	mus musculus	227	6	2.8	632	1	GAAT_HUMAN	Q9un88	homo sapien
155	6	2.8	413	1	ARCA_CLOPE	Q46254	clostridium	228	6	2.8	632	1	GIDA_BACHD	Q9K6c1	baecillus ha
156	6	2.8	413	1	YAA7_SCHPO	O09800	schizosacch	229	6	2.8	633	1	CYSN_PSEAE	O50274	p cyen/cybc
157	6	2.8	414	1	PIAT_MYCLE	O50173	mycobacteri	230	6	2.8	637	1	YHE7_YEAST	P38724	saccharomyc
158	6	2.8	417	1	HS47_HUMAN	P29943	homo sapien	231	6	2.8	640	1	RAEP_CANAL	O33831	candida alb
159	6	2.8	417	1	HS47_MOUSE	P19324	mus musculus	232	6	2.8	642	1	Y202_AERPE	O9Y916	aeropyrum p
160	6	2.8	417	1	HS47_RAT	P29457	rattus norv	233	6	2.8	644	1	NEM_RABIT	P54938	oryctolagus
161	6	2.8	417	1	PIT_MYCTU	O06411	mycobacteri	234	6	2.8	653	1	RAE1_HUMAN	P23386	homo sapien
162	6	2.8	418	1	CBP2_HUMAN	P50454	homo sapien	235	6	2.8	655	1	ILV1_BRANA	P27818	brasica na
163	6	2.8	423	1	ENO_MERVA	Q60173	methanococc	236	6	2.8	657	1	GRAD_TREPA	O83062	treponema p
164	6	2.8	441	1	HISX_STRCO	P16455	strepococc	237	6	2.8	662	1	ACSA_HELPY	O35686	helicobacte
165	6	2.8	445	1	EMA7_YEAST	P43555	saccharomyc	238	6	2.8	663	1	MX1_PIG	P27390	sus scrofa
166	6	2.8	446	1	ATPD_MYCLE	P53006	mycobacteri	239	6	2.8	671	1	CHEA_THEMA	O56310	thermotoga
167	6	2.8	446	1	ATPD_MOUSE	Q64163	mus musculus	240	6	2.8	672	1	STGB_THEMA	O9Y660	thermotoga
168	6	2.8	447	1	DHB4_CORGL	Q64434	mus musculus	241	6	2.8	676	1	VGP_EBOSH	Q68814	ebola virus
169	6	2.8	447	1	PTK6_MOUSE	P31026	corynebacte	242	6	2.8	676	1	VGP_EBOSH	Q68798	ebola virus
170	6	2.8	452	1	TYR1_YEAST	P20049	azorhizobiu	243	6	2.8	677	1	WHIT_LYCCU	Q05360	jucilla cup
171	6	2.8	452	1	NTRY_AZOCA	Q04449	azorhizobiu	244	6	2.8	678	1	VF48_MYCTU	O10778	mycobacteri
172	6	2.8	457	1	EMB8_PICGL	Q40863	picea glauc	245	6	2.8	679	1	WHIT_CERCA	O17320	ceratitlis c
173	6	2.8	461	1	USP_MANSE	P41758	chlamydomon	246	6	2.8	680	1	APB2_MOUSE	P98084	mus musculus
174	6	2.8	461	1	USP_MANSE	P54779	manduca sex	247	6	2.8	681	1	CAO2_RABIT	O02767	o acyl-coen
175	6	2.8	463	1	PLSB_CARTI	Q42713	carthamus t	248	6	2.8	683	1	ACSI_KLUDA	O60011	kluyveromyc
176	6	2.8	464	1	SMAS_BRARE	O9Y7e7	brachydanio	249	6	2.8	684	1	SGT1_DROME	O90403	drosophila
177	6	2.8	465	1	SMAL_HUMAN	O15797	homo sapien	250	6	2.8	685	1	YYCA_BACSU	P37483	baecillus su
178	6	2.8	465	1	SMAL_MOUSE	P70340	mus musculus	251	6	2.8	687	1	WHIT_DROME	P10090	drosophila
179	6	2.8	465	1	SMAS_HUMAN	Q99717	homo sapien	252	6	2.8	689	1	YUXG_BACSU	P40747	baecillus su

253	6	2.8	695	1	PID1_BOVIN	P10895	bov taurus	326	6	2.8	1586	1	SN22_HUMAN	P51511	homo sapien
254	6	2.8	733	1	TKT_FICST	P34756	pichia atip	327	6	2.8	1587	1	SUR2_CAEEL	Q10669	caenorhabdit
255	6	2.8	619	1	ACSI1_YEAST	O01574	saccharomyc	328	6	2.8	1634	1	PK350_HUMAN	O00750	homo sapien
256	6	2.8	716	1	GFAL1_YEAST	P14742	saccharomyc	329	6	2.8	1720	1	FTSH_CHLYU	P56319	chlamydia v
257	6	2.8	734	1	VTBR_HAYEB	P28969	equine hearp	330	6	2.8	1836	1	UVRA_CHLPP	O92985	chlamydia p
258	6	2.8	738	1	YOLJ1_BP42	P42548	bacterioph	331	6	2.8	1972	1	MYHB_HUMAN	P35749	homo sapien
259	6	2.8	745	1	PEC1_PIG	Q95242	sus scrofa	332	6	2.8	1972	1	MYHB_MOUSE	O08638	mus musculus
260	6	2.8	745	1	PNP_RICPR	Q92d43	rickettsia	333	6	2.8	1972	1	MYHB_RABIT	P07948	oryctolagus
261	6	2.8	746	1	FEPA_ECOLI	P05825	escherichia	334	6	2.8	2109	1	PGCA_CHICK	P07888	gallus gall
262	6	2.8	749	1	APB2_HUMAN	O99767	homo sapien	335	6	2.8	2265	1	PINC_BOVIN	P07558	bos taurus
263	6	2.8	750	1	APB2_RAT	O35431	rattus norv	336	6	2.8	2303	1	POLG_TMEVD	P13889	genome po
264	6	2.8	752	1	CO2_HUMAN	P06681	homo sapien	337	6	2.8	2303	1	POLG_TMEVB	P08544	genome po
265	6	2.8	756	1	PID1_HUMAN	P51178	homo sapien	338	6	2.8	2349	1	TPR_HUMAN	P12270	homo sapien
266	6	2.8	756	1	PID1_RAT	P10688	rattus norv	339	6	2.8	2358	1	YEBJ_ECOLI	P76347	escherichia
267	6	2.8	760	1	CO2_MOUSE	P21180	mus musculus	340	6	2.8	2415	1	SPCA_DROME	P13345	drosophila
268	6	2.8	778	1	RIR1_ASFB7	P42491	african swi	341	6	2.8	2663	1	CENE_HUMAN	O03224	homo sapien
269	6	2.8	779	1	RIR1_ASFB2	P26685	african swi	342	6	2.8	2688	1	ZEP1_MOUSE	O01122	mus musculus
270	6	2.8	799	1	ITBN_DROME	Q27591	drosophila	343	6	2.8	2717	1	ZEP1_HUMAN	P15822	homo sapien
271	6	2.8	805	1	AHR_MOUSE	P30561	mus musculus	344	6	2.8	2834	1	IP3R_DROME	P27993	drosophila
272	6	2.8	810	1	NFM_BOVIN	O77788	bos taurus	345	6	2.8	3011	1	POLG_HCVH	P27958	genome po
273	6	2.8	821	1	MCM6_MOUSE	P97311	mus musculus	346	6	2.8	3301	1	CLR3_MOUSE	O91430	mus musculus
274	6	2.8	825	1	GUN3_BACS4	P19570	bacillus sp	347	6	2.8	3312	1	CLR3_HUMAN	O9HY57	homo sapien
275	6	2.8	831	1	NPH_RAT	P16884	rattus norv	348	6	2.8	3313	1	CLR3_RAT	O88278	rattus norv
276	6	2.8	837	1	APB1_HUMAN	Q02410	homo sapien	349	6	2.8	3367	1	ERY2_SACER	O03132	saccharopol
277	6	2.8	839	1	APB1_RAT	O35430	rattus norv	350	6	2.8	3951	1	VGPI_IBVB	P27920	avian infec
278	6	2.8	840	1	GYRA_UREPA	Q9P163	ureaplasma	351	6	2.8	4128	1	PRKD_HUMAN	P78527	homo sapien
279	6	2.8	845	1	NFM_RAT	P12839	rattus norv	352	6	2.8	4128	1	EMF1_ECOLI	P20860	escherichia
280	6	2.8	845	1	PPSA_AERPE	O9Y655	aeropyrum p	353	5	2.3	23	1	CH60_THIFE	P29134	thiobacillus
281	6	2.8	848	1	NFM_MOUSE	P08553	mus musculus	354	5	2.3	23	1	CAMT_PIRNS	P81081	pinus pinas
282	6	2.8	880	1	CADF_XENLA	P33148	xenopus lae	355	5	2.3	38	1	PHRG_BACSU	O32255	bacillus su
283	6	2.8	886	1	APCE_PORPU	P31263	porphyra pu	356	5	2.3	39	1	PHRI_BACSU	O31429	bacillus su
284	6	2.8	886	1	SUHW_DROAN	Q08875	drosophila	357	5	2.3	40	1	CH60_BACFO	P81284	bacteroides
285	6	2.8	886	1	STU2_YEAST	P46675	saccharomyc	358	5	2.3	43	1	VG67_BPM45	Q05221	mycoplacte
286	6	2.8	894	1	LOL1_DROME	P42683	drosophila	359	5	2.3	44	1	LPAS_STRMU	P10540	streptococc
287	6	2.8	903	1	ECIC_BOVIN	P54281	bos taurus	360	5	2.3	51	1	IP21_SOLTU	P01079	solanum tub
288	6	2.8	904	1	SYA_EUTRO	Q97114	bulfiolobus	361	5	2.3	51	1	STC3_SUTSO	P81551	bulfiolobus
289	6	2.8	911	1	CAIB_BOVIN	Q28083	bos taurus	362	5	2.3	56	1	SSPF_BACME	P52291	bacillus me
290	6	2.8	920	1	MMU7_MYCTU	P96289	mycobacteri	363	5	2.3	57	1	RIC1_PXYIN	Q9Y068	phytophthor
291	6	2.8	921	1	ITH4_PIG	P79263	sus scrofa	364	5	2.3	59	1	SSPF_BACCE	P52290	bacillus ce
292	6	2.8	925	1	N107_HUMAN	P57149	homo sapien	365	5	2.3	63	1	RPO9_MYXVL	Q9G886	myxoma viru
293	6	2.8	926	1	N107_RAT	P52590	rattus norv	366	5	2.3	64	1	Y160_BP74	P39321	bacterioph
294	6	2.8	929	1	CA1C_NOTVI	Q91145	notophthalm	367	5	2.3	64	1	Y7K5_VACCV	P18353	vaccinia vi
295	6	2.8	930	1	ITH4_HUMAN	Q14624	h inter-alp	368	5	2.3	67	1	RS21_AOUAE	O67028	aquifex aeo
296	6	2.8	937	1	ODO1_BACSU	P31329	bacillus su	369	5	2.3	67	1	Y03G_BP74	P13337	bacterioph
297	6	2.8	955	1	AGAB_VIBS7	P48840	vibrio sp.	370	5	2.3	69	1	YDHZ_ECOLI	P77274	escherichia
298	6	2.8	962	1	AMPN_PIG	P15145	sus scrofa	371	5	2.3	70	1	V43_BP73	P20317	bacterioph
299	6	2.8	965	1	AMPN_RABIT	P15541	oryctolagus	372	5	2.3	70	1	YDHZ_SALTY	Q92451	salmonella
300	6	2.8	966	1	AMPN_HUMAN	P15144	homo sapien	373	5	2.3	71	1	Y1K6_BP71	P10430	bacterioph
301	6	2.8	967	1	SYA_BOMMO	P21894	bombyx mori	374	5	2.3	71	1	RLJ3_YERPE	P58471	yersinia pe
302	6	2.8	969	1	SACB_STRSL	Q55242	streptococc	375	5	2.3	72	1	VX2A_CVPPU	P09046	porcine tra
303	6	2.8	988	1	PKC1_SCHPO	P36582	schistosom	376	5	2.3	72	1	MTBG_METMA	P80656	methanogarc
304	6	2.8	1007	1	Y741_CHLMU	O9P166	chlamydia m	377	5	2.3	74	1	YDPC_ECOLI	P1418	escherichia
305	6	2.8	1034	1	BGAL_BACME	Q52847	bacillus me	378	5	2.3	74	1	COX3_DROSI	P50271	drosophila
306	6	2.8	1046	1	SBCC_LACLA	Q9C280	lactococcus	379	5	2.3	74	1	SECG_UREPA	Q9P159	ureaplasma
307	6	2.8	1077	1	CTR9_YEAST	P89105	saccharomyc	380	5	2.3	75	1	YF67_THEMA	Q9X143	thermoga
308	6	2.8	1086	1	POL_OYVVS	P16901	cvine lenti	381	5	2.3	76	1	RPOZ_AOUAE	O66570	aquifex aeo
309	6	2.8	1094	1	A3B1_HUMAN	O00203	homo sapien	382	5	2.3	76	1	RS37_YEAST	P05759	saccharomyc
310	6	2.8	1119	1	RPOB_THEAO	Q9KWT7	thermus aqu	383	5	2.3	76	1	SECG_MYCPN	O96X08	mycoplasma
311	6	2.8	1121	1	EXSC_HARIN	P44945	haemophilus	384	5	2.3	77	1	PUP0_RHOSH	P16069	rhodobacter
312	6	2.8	1146	1	ITAI_DROME	Q24247	drosophila	385	5	2.3	77	1	SECG_CONGL	Q92469	corynebacte
313	6	2.8	1148	1	A8A2_MOUSE	P98200	mus musculus	386	5	2.3	77	1	SECG_MYCGE	P58061	mycoplasma
314	6	2.8	1191	1	N1FJ_RHORU	Q63046	rhodospirill	387	5	2.3	77	1	TATI_AOUAE	O66476	aquifex aeo
315	6	2.8	1192	1	RTN4_HUMAN	O9NG36	homo sapien	388	5	2.3	78	1	Y13C_BP74	P71307	bacterioph
316	6	2.8	1195	1	YK76_YEAST	P36168	saccharomyc	389	5	2.3	80	1	PYS1_ANASP	P07121	anabaena sp
317	6	2.8	1337	1	PRPJ_HUMAN	Q12913	homo sapien	390	5	2.3	81	1	S112_RABIT	O77721	oryctolagus
318	6	2.8	1355	1	DP3A_SYNY3	P74750	bynechocyst	391	5	2.3	81	1	ULI1_VZVD	P09257	varicella-z
319	6	2.8	1362	1	CA21_CHICK	P02467	gallus gall	392	5	2.3	81	1	YP95_MYCTU	O50626	mycobacteri
320	6	2.8	1371	1	VCAP_HSVCA	O00959	herpesviru	393	5	2.3	84	1	YVDA_VACCV	P04229	vaccinia vi
321	6	2.8	1372	1	CA21_MOUSE	O01149	mus musculus	394	5	2.3	86	1	OXIR_PIG	O97661	sus scrofa
322	6	2.8	1442	1	DPO3_UREPA	Q9P044	ureaplasma	395	5	2.3	87	1	VP08_BRAPS	O9TUD0	bacterioph
323	6	2.8	1481	1	RPOD_ODOSI	P49466	odontella s	396	5	2.3	87	1	Y1K6_BP71	P10435	bacterioph
324	6	2.8	1505	1	DPOA_DROME	P26019	drosophila	397	5	2.3	88	1	PTHP_MYCCA	P45611	mycoplasma
325	6	2.8	1556	1	GLTS_SYNY3	P55038	bynechocyst	398	5	2.3	88	1	R25Z_SCHPO	O74172	schistosach

399	5	2.3	88	1	VG15_BPMU	09L1x4 bacterioph	472	5	2.3	113	1	MEAL_PIG	09S313 sus scrofa
400	5	2.3	89	1	DSBA_BPT4	P13320 bacterioph	473	5	2.3	113	1	RBS3_WHEAT	P07398 triticum ae
401	5	2.3	89	1	R25Y_SCHPO	P70009 schizosacch	474	5	2.3	114	1	RNU2_USRSP	P00654 ussilaago sp
402	5	2.3	89	1	SV16_HUMAN	P557774 h small ind	475	5	2.3	114	1	Y116_ARCFU	030118 archaeoglob
403	5	2.3	89	1	YTA3_SULSO	P58017 sulfolobus	476	5	2.3	114	1	Y074_AOUAE	067784 aquifex aeo
404	5	2.3	90	1	RS16_LACIA	Q9C1b2 lactococcus	477	5	2.3	114	1	YNS8_CABEL	P34552 caenorhabdi
405	5	2.3	90	1	VPM_BBPRD	P27389 bacterioph	478	5	2.3	115	1	P1XT_RHIME	086454 rhizobium m
406	5	2.3	91	1	VE4_HPV58	P26549 human papil	479	5	2.3	115	1	P722_STYPL	P28214 styeila plic
407	5	2.3	93	1	ACYF_MYCTU	P56543 mycobacteri	480	5	2.3	115	1	PTHY_RAT	P04089 rattus norv
408	5	2.3	94	1	RL35_PIG	Q29361 sus scrofa	481	5	2.3	115	1	VAPZ_BACNO	Q46511 bacteroides
409	5	2.3	94	1	YNOF_HALMA	P22527 haloarcula	482	5	2.3	115	1	YPOL_IDBVP	P25222 avian infec
410	5	2.3	95	1	Y229_METUA	Q57682 methanococc	483	5	2.3	116	1	CH15_DROVI	P133424 drosophila
411	5	2.3	95	1	YPS8_RHILP	P10497 rhizobium 1	484	5	2.3	117	1	HV58_MOUSE	P18529 mus musculu
412	5	2.3	96	1	MYNA_MYTGA	P82103 mytilus gal	485	5	2.3	117	1	YXEF_BACPA	Q45345 chlamydia pa
413	5	2.3	97	1	SFT1_YEAST	P43682 saccharomyc	486	5	2.3	117	1	YCX2_CHIRE	P05722 chlamydomon
414	5	2.3	97	1	YAN9_YEAST	P39565 saccharomyc	487	5	2.3	118	1	RL17_AOUAE	066482 aquifex aeo
415	5	2.3	98	1	ELI6_PHYCT	P15565 phytophthor	488	5	2.3	119	1	ACPS_MYCPN	P75480 mycoplasma
416	5	2.3	98	1	NDLM_LUMTE	Q34948 lumbricus t	489	5	2.3	119	1	MOSA_CLOPE	Q6X1n2 clostridium
417	5	2.3	98	1	S113_BOVIN	P79342 bos taurus	490	5	2.3	119	1	Y836_METUA	Q58246 methanococc
418	5	2.3	98	1	S113_HUMAN	Q99584 homo sapien	491	5	2.3	119	1	Y033_ARCFU	030238 archaeoglob
419	5	2.3	98	1	S113_MOUSE	P97352 mus musculu	492	5	2.3	120	1	GLH1_RAT	P11962 r glycoprot
420	5	2.3	99	1	GP45_BPSP1	Q48399 bacterioph	493	5	2.3	120	1	GLH2_RAT	P01217 b glycoprot
421	5	2.3	99	1	HRO_ERWCA	Q92176 erwinia car	494	5	2.3	120	1	GLHA_BOVIN	Q9X848 c glycoprot
422	5	2.3	99	1	KAP3_MOUSE	P31324 mus musculu	495	5	2.3	120	1	GLHA_CANPA	Q28345 e glycoprot
423	5	2.3	99	1	TRBD_RHISN	P55397 rhizobium s	496	5	2.3	120	1	GLHA_EQUAS	Q28345 e glycoprot
424	5	2.3	99	1	Y754_AOUAE	Q66956 aquifex aeo	497	5	2.3	120	1	GLHA_HORSE	P01220 e glycoprot
425	5	2.3	99	1	YHCO_BACSU	P54599 bacillus su	498	5	2.3	120	1	GLHA_MACMU	Q46687 m glycoprot
426	5	2.3	99	1	YHCO_BACSU	P54599 bacillus su	499	5	2.3	120	1	GLHA_MACRO	Q46687 m glycoprot
427	5	2.3	100	1	KOAI_ECOLI	Q57423 escherichia	500	5	2.3	120	1	GLHA_MASCO	Q46687 m glycoprot
428	5	2.3	100	1	YGPE_HAEIN	P44062 haemophilus	501	5	2.3	120	1	GLHA_MERUN	Q46687 m glycoprot
429	5	2.3	101	1	KOAI_ECOLI	P03052 escherichia	502	5	2.3	120	1	GLHA_MESAU	Q46687 m glycoprot
430	5	2.3	101	1	RS16_UREPA	Q90981 ureaplasma	503	5	2.3	120	1	GLHA_MOUSE	P01216 m glycoprot
431	5	2.3	101	1	VNS7_CVPE3	P33465 feline ente	504	5	2.3	120	1	GLHA_PIG	P01216 m glycoprot
432	5	2.3	101	1	VNS7_F1PV	P19742 feline infe	505	5	2.3	120	1	GHP_HORSE	P02726 equus cabal
433	5	2.3	102	1	HG14_CHICK	P12274 gallus gall	506	5	2.3	120	1	RTL3_MARKO	P26872 marcanthia
434	5	2.3	102	1	RL12_METUA	P54048 methanococc	507	5	2.3	120	1	VGB_BP4	P03634 bacterioph
435	5	2.3	102	1	RS24_HALNI	Q9Hn14 halobacteri	508	5	2.3	120	1	CH15_DROSV	P133426 drosophila
436	5	2.3	102	1	SRI9_SULSO	Q980W2 sulfolobus	509	5	2.3	121	1	GAT4_YEAST	CH15_DROSV
437	5	2.3	103	1	PTLA_STAMU	Q02809 staphylococc	510	5	2.3	121	1	PEDB_MERTH	P04056 saccharomyc
438	5	2.3	104	1	PTLA_STAMU	P26426 streptococc	511	5	2.3	121	1	RA14_CANAL	PE3709 candida alb
439	5	2.3	104	1	VU9_HAVBU	Q01347 human heipe	512	5	2.3	121	1	VOR4_P1AMV	Q07520 plantago as
440	5	2.3	105	1	GLRX_HUMAN	P35754 homo sapien	513	5	2.3	121	1	YCM1_YEAST	P25642 saccharomyc
441	5	2.3	105	1	YNT7_CAUCR	Q9A580 caulobacter	514	5	2.3	121	1	ACPS_FUSNN	Q46747 fusobacteri
442	5	2.3	105	1	YNC8_YEAST	P53967 saccharomyc	515	5	2.3	122	1	RL14_HELPJ	Q9Z188 helicobacte
443	5	2.3	105	1	APV1_CHICK	P94500 bacillus su	516	5	2.3	122	1	RL14_HELPJ	Q9Z188 helicobacte
444	5	2.3	106	1	APV1_CHICK	P02659 gallus gall	517	5	2.3	122	1	RL14_HELPJ	Q9Z188 helicobacte
445	5	2.3	106	1	GLRX_MOUSE	Q9qnh0 mus musculu	518	5	2.3	122	1	RL14_HELPJ	P33100 micrococcu
446	5	2.3	106	1	GLRX_RAT	Q9esh6 rattus norv	519	5	2.3	122	1	RL14_HELPJ	P42766 homo sapien
447	5	2.3	106	1	IDHG_PIG	P41566 sus scrofa	520	5	2.3	122	1	RL14_HELPJ	P17078 rattus norv
448	5	2.3	106	1	YB43_VIBCH	Q9K8W3 vibrio chol	521	5	2.3	122	1	RL14_HELPJ	Q46912 guillardi
449	5	2.3	106	1	YK53_YEAST	P36153 saccharomyc	522	5	2.3	122	1	VNS2_INBYA	P46051 anabaena va
450	5	2.3	107	1	RLA2_TRYBB	P51408 trypanosoma	523	5	2.3	123	1	HEB1_ANAVA	P18501 anabaena va
451	5	2.3	107	1	UI19_ARATH	Q98888 arabidopsis	524	5	2.3	123	1	HEB1_ANASP	P00870 spinacia oi
452	5	2.3	107	1	YFHP_ECOLI	P36539 escherichia	525	5	2.3	123	1	RBS1_SPTOL	P00870 spinacia oi
453	5	2.3	108	1	PRV8_CYPCA	P02618 cyprinus ca	526	5	2.3	123	1	RL35_CABEL	P34662 caenorhabdi
454	5	2.3	108	1	RS25_YEAST	Q9vri8 drosophila	527	5	2.3	124	1	OLPX_CHLTR	Q44673 chlamydia t
455	5	2.3	108	1	UI19_DROME	Q9vri8 drosophila	528	5	2.3	124	1	VCL1_PEA	P02856 pisum sativ
456	5	2.3	109	1	K2M1_SHEEP	Q02539 ovis aries	529	5	2.3	125	1	GLGS_MAIZE	P55240 zea mays (m
457	5	2.3	109	1	YQCC_ECOLI	Q46919 escherichia	530	5	2.3	125	1	SPIS_COMTE	P00947 comamonas t
458	5	2.3	110	1	KDGF_ERWCH	Q05527 erwinia chr	531	5	2.3	126	1	PRD4_AVERA	Q9M464 avena fatua
459	5	2.3	110	1	KDGF_ERWCH	Q05527 erwinia chr	532	5	2.3	126	1	PRD4_AVERA	P25554 caenorhabdi
460	5	2.3	110	1	YCI3_METUA	P01694 eryctolagus	533	5	2.3	126	1	RS13_AOUAE	Q64486 aquifex aeo
461	5	2.3	111	1	YCI3_METUA	P01694 eryctolagus	534	5	2.3	127	1	CRCB_THEMEA	Q9XW88 thermotoga
462	5	2.3	111	1	YR2C_BACSU	Q34527 bacillus su	535	5	2.3	127	1	PAND_NEIMA	Q91256 neisseria m
463	5	2.3	112	1	CL34_DROMI	Q01774 drosophila	536	5	2.3	127	1	PAND_NEIMA	Q91256 neisseria m
464	5	2.3	112	1	CLP4_DROME	P01789 drosophila	537	5	2.3	127	1	PK12_ODOSI	P01789 drosophila
465	5	2.3	112	1	FER1_CAUCR	Q45572 caulobacter	538	5	2.3	127	1	RL7_CAUCR	Q9A847 caenorhabdi
466	5	2.3	112	1	GLNB_AZOB	P21193 azospirillum	539	5	2.3	128	1	SYR8_HUMAN	Q9A847 caenorhabdi
467	5	2.3	112	1	OL11_MOUSE	Q60890 mus musculu	540	5	2.3	128	1	Y14D_BP4	P33507 bacterioph
468	5	2.3	112	1	OL11_MOUSE	Q60890 mus musculu	541	5	2.3	128	1	Y14D_BP6	Q38169 bacterioph
469	5	2.3	112	1	RR6_PORPU	P51359 porphyra pu	542	5	2.3	128	1	Y280_METUA	Q57728 methanococc
470	5	2.3	112	1	THL_CLODI	P45362 clostridium	543	5	2.3	128	1	Y532_BUCAI	P57598 buchnera ap
471	5	2.3	113	1	HYPA_RHOCA	P26409 rhodobacter	544	5	2.3	128	1	Y612_ARCFU	029643 archaeoglob

545	5	2.3	128	1	YCDK_ECOLI	P75896	eecherichia	618	5	2.3	146	1	RL15_BACST	P04452	bacillus st
546	5	2.3	128	1	YN31_MYCTU	P71881	mycobacteri	619	5	2.3	147	1	DTD_STREPN	Q973p3	streplococc
547	5	2.3	128	1	YPRR_ECOLI	P03848	eecherichia	620	5	2.3	147	1	GLPI_GLYDI	P332t1	glycera dib
548	5	2.3	130	1	CM36_YEAST	P25603	saccharomyc	621	5	2.3	147	1	GLP2_GLYDI	P21659	glycera dib
549	5	2.3	130	1	SV28_MOUSE	Q9j1l2	mus musculus	622	5	2.3	147	1	GS17_XENLA	P07733	xenopus lae
550	5	2.3	130	1	TV41_MOUSE	P01738	mus musculus	623	5	2.3	147	1	INNH_SALCH	P37422	salmonella
551	5	2.3	131	1	AGSW_MOUSE	O03288	mus musculus	624	5	2.3	147	1	INNH_SALCH	P37422	salmonella
552	5	2.3	131	1	AGSW_MOUSE	P79407	vulpes vulp	625	5	2.3	147	1	YHBP_ECOLI	P54711	eecherichia
553	5	2.3	131	1	FABP_LEPDS	Q9u520	lepidoglyph	626	5	2.3	147	1	YHBP_ECOLI	P54711	eecherichia
554	5	2.3	131	1	NUSB_BACSU	P54520	bacillus su	627	5	2.3	148	1	FUR_YERPE	P33086	mycobacteri
555	5	2.3	131	1	RK12_CHLVU	P56345	chlorella v	628	5	2.3	148	1	GCSS_AOUAB	O66720	aquifex ae
556	5	2.3	132	1	AGSW_HUMAN	P42127	homo sapien	629	5	2.3	148	1	PUIA_WHEAT	P33432	trititum ae
557	5	2.3	133	1	AGSW_BOVIN	Q29414	bos taurus	630	5	2.3	148	1	RL2A_EUPCR	P81421	euplotes cr
558	5	2.3	133	1	CDI_MYCPI	P47718	mycoplasma	631	5	2.3	148	1	SSRP_FUSNP	Q9rfe8	haemophilus
559	5	2.3	133	1	S3AD_BACSU	P19739	human coron	632	5	2.3	148	1	Y246_HAEIN	P43792	haemophilus
560	5	2.3	133	1	VNS4_CV422	P19739	human coron	633	5	2.3	148	1	Y246_HAEIN	P43792	haemophilus
561	5	2.3	133	1	Y014_BRHP1	P51716	bacterioph	634	5	2.3	148	1	YF56_MYCTU	Q10772	mycobacteri
562	5	2.3	134	1	NUSB_BACHD	Q9K965	bacillus ha	635	5	2.3	149	1	RL9_BACST	P02417	bacillus st
563	5	2.3	134	1	TRBC_AGRU	P54908	agrobacteri	636	5	2.3	149	1	RL9_HAEIN	P44349	haemophilus
564	5	2.3	134	1	WNT1_CHICK	Q91029	gallus gall	637	5	2.3	150	1	PLSC_PHYPO	P19203	phyasium po
565	5	2.3	135	1	FKB2_YEAST	P32472	saccharomyc	638	5	2.3	150	1	REG9_PYRAB	Q9uyl6	pyrococcus
566	5	2.3	135	1	RL27_ARATH	P51419	arabidopsis	639	5	2.3	150	1	RL9_STREPY	O99x16	streplococc
567	5	2.3	135	1	RL27_PEA	Q05462	pisum sativ	640	5	2.3	150	1	SSRP_BORBU	O31064	botreilia bu
568	5	2.3	135	1	RS6_ECOLI	P02358	eecherichia	641	5	2.3	150	1	YPSL_SYNEH	P25968	synechococc
569	5	2.3	135	1	TV43_HUMAN	P01737	homo sapien	642	5	2.3	151	1	DUT_HAEIN	P43792	haemophilus
570	5	2.3	136	1	EF1G_PIG	Q29387	sus scrofa	643	5	2.3	151	1	DUT_PASMU	P57914	pasturella
571	5	2.3	136	1	MUTT_HAEIN	P44932	haemophilus	644	5	2.3	151	1	DUT_YERPE	O8zjf5	yersinia pe
572	5	2.3	136	1	YAI3_ARCPO	O29249	archaeoglob	645	5	2.3	151	1	HS11_WHEAT	P12810	trititum ae
573	5	2.3	136	1	YAI3_ARCPO	O29249	archaeoglob	646	5	2.3	152	1	ANP_MOUSE	P05151	mus musculus
574	5	2.3	137	1	INL2_DROME	Q9vct51	drosophila	647	5	2.3	152	1	MR4Z_ECO57	P01161	rattus norv
575	5	2.3	137	1	MALI_VIBRU	P96158	vibrio furn	648	5	2.3	152	1	MR4Z_ECO57	Q8x923	eecherichia
576	5	2.3	137	1	MBA2_ECOLI	P07113	eecherichia	649	5	2.3	152	1	MR4Z_ECOLI	P22186	eecherichia
577	5	2.3	138	1	HUPG_RHOCA	O03006	rhodobacter	650	5	2.3	152	1	MR4Z_SALTI	O8z295	salmonella
578	5	2.3	138	1	NUSB_YERPE	O8zc42	yersinia pe	651	5	2.3	152	1	RECK_VIBCH	O8z295	salmonella
579	5	2.3	138	1	RL27_SOLTU	P41101	solanum tub	652	5	2.3	152	1	RECK_VIBCH	Q66647	vibrio chol
580	5	2.3	138	1	YUC7_YEAST	P47063	saccharomyc	653	5	2.3	153	1	RBM3_MOUSE	Q89086	mus musculus
581	5	2.3	139	1	GRA1_ORYSA	Q07077	oryza sativ	654	5	2.3	153	1	HS11_LYCSE	P30121	lycopersico
582	5	2.3	139	1	LEOV_MYCTU	O53412	mycobacteri	655	5	2.3	154	1	NRDG_ECOLI	P39339	eecherichia
583	5	2.3	139	1	NUSB_ECOLI	P04381	eecherichia	656	5	2.3	154	1	NRDG_ECOLI	O8z138	salmonella
584	5	2.3	139	1	NUSB_SALTI	O8z8x6	salmonella	657	5	2.3	154	1	NRDG_SALTI	Q91645	salmonella
585	5	2.3	139	1	NUSB_SALTY	O8z8x6	salmonella	658	5	2.3	154	1	PPDA_AERPE	Q9y428	aeropyrum p
586	5	2.3	139	1	R1SB_METTH	O27443	methanobact	659	5	2.3	154	1	SSRP_SYNY3	P74355	synechocyst
587	5	2.3	140	1	HBAD_COLLI	O12985	columba liv	660	5	2.3	154	1	YAI6_ARCFU	O29246	archaeoglob
588	5	2.3	140	1	NUSB_STRPN	O97ef0	streplococc	661	5	2.3	154	1	YD88_SYNY3	P74148	synechocyst
589	5	2.3	140	1	VG07_HSV1	O00133	ictaluriid h	662	5	2.3	154	1	YK01_CAREL	P34429	caenorhabdi
590	5	2.3	141	1	GAT3_YEAST	O07928	saccharomyc	663	5	2.3	154	1	YK01_CAREL	Q50665	mycobacteri
591	5	2.3	141	1	HBAD_PHRHI	P02006	phrynos h	664	5	2.3	155	1	BCT1_MYCTU	P22226	bos taurus
592	5	2.3	141	1	MMOB_METCA	P18797	methylococc	665	5	2.3	155	1	BCT1_SHEEP	P54230	ovis aries
593	5	2.3	141	1	SP52_BACSU	P40867	bacillus su	666	5	2.3	155	1	GP27_BPSP1	P66228	bacterioph
594	5	2.3	141	1	SSRP_UREPA	O99x87	ureaplasma	667	5	2.3	155	1	NRDG_VIBCH	Q9kmt6	vibrio chol
595	5	2.3	142	1	FUSB_BURCE	P24127	burkholderi	668	5	2.3	155	1	SSRP_ANASP	O8ymt0	anabena sp
596	5	2.3	143	1	HS11_MEDSA	P27879	medicago ba	669	5	2.3	155	1	SSRP_CLOAB	Q97149	clostridium
597	5	2.3	143	1	IR09_HCVNA	P16807	human cytom	670	5	2.3	155	1	YM27_XYLFA	Q9p0b7	xylotella fas
598	5	2.3	143	1	NUSB_BUCAI	P57535	buchnera ap	671	5	2.3	156	1	HS11_ARATH	P19056	arabidopsis
599	5	2.3	143	1	PSE2_NICSY	O41229	nicotiana s	672	5	2.3	156	1	NRDL_STRPN	Q97fc3	streplococc
600	5	2.3	144	1	DTD_PASMU	O9ckk0	pasturella	673	5	2.3	156	1	NUSB_RICCN	Q92f65	rickettsia
601	5	2.3	144	1	DTD_VIBCH	O9enr7	vibrio chol	674	5	2.3	156	1	NUSB_VIBCH	O9knp5	vibrio chol
602	5	2.3	144	1	LRP5_BACSU	P65582	bacillus su	675	5	2.3	156	1	RL22_MERTU	P54035	methanococc
603	5	2.3	144	1	RL22_MYCGE	P47402	mycoplasma	676	5	2.3	156	1	RS7_MITCU	P94088	micrococc
604	5	2.3	145	1	ANP_RANCA	P18909	rana catesb	677	5	2.3	157	1	SSRP_CLOPE	Q8xk12	clostridium
605	5	2.3	145	1	DTD_ECOLI	P32147	eecherichia	678	5	2.3	157	1	ATPX_GALSU	P35012	galidieria s
606	5	2.3	145	1	DTD_SALTI	P56533	salmonella	679	5	2.3	157	1	HS12_ARATH	P13863	arabidopsis
607	5	2.3	145	1	DTD_SALTY	P56533	salmonella	680	5	2.3	157	1	ISPF_BACHD	Q9kfe3	bacillidops
608	5	2.3	145	1	DTD_YERPE	P56534	yersinia pe	681	5	2.3	157	1	MANB_KLEPN	Q846c3	klebsiella
609	5	2.3	145	1	DUT_CLOAB	Q97j61	clostridium	682	5	2.3	157	1	RUVK_HELPJ	Q9zxc3	helicobacte
610	5	2.3	145	1	OM25_HUMAN	P57105	homo sapien	683	5	2.3	157	1	RUVK_HELPJ	Q9zxc3	helicobacte
611	5	2.3	145	1	PSE2_ARATH	O98716	arabidopsis	684	5	2.3	157	1	SSB_CHLTR	O25544	helicobacte
612	5	2.3	145	1	RM11_RECAM	O21234	reclinomona	685	5	2.3	157	1	YMA7_CAREL	O84408	chlamydia t
613	5	2.3	145	1	SSRP_MYCPU	O986q6	mycoplasma	686	5	2.3	158	1	FLAV_BACSU	P43737	bacillus su
614	5	2.3	145	1	YPOL_IBDVS	P25221	avian infec	687	5	2.3	158	1	GRBB_ECOLI	P30128	eecherichia
615	5	2.3	145	1	YPOL_IBDVS	P15481	avian infec	688	5	2.3	158	1	HS11_PEA	P19243	pisum sativ
616	5	2.3	145	1	YPOL_IBDVS	P22440	avian infec	689	5	2.3	158	1	PGSG_HUMAN	P10124	homo sapien
617	5	2.3	146	1	HBB_ERIBU	P02059	eritaceus e	690	5	2.3	158	1	VH21_MYXVL	P28850	myxoma viru

691	5	2.3	158	1	XGPT_BUCAI	P57339	buchnera ap	764	5	2.3	172	1	HSLV_PASHA	P49617	pasteurella
692	5	2.3	158	1	XKFG_ECOLI	Q47685	escherichia	765	5	2.3	172	1	MEAI_HUMAN	Q16626	homo sapien
693	5	2.3	159	1	BYKE_BETVE	P43178	betula verr	766	5	2.3	172	1	MTHG_DROME	Q5877	drosophila
694	5	2.3	159	1	BVJF_BETVE	P43179	betula verr	767	5	2.3	172	1	THIM_ORYSA	Q92P20	oryza sativ
695	5	2.3	159	1	BVJL_BETVE	P43183	betula verr	768	5	2.3	172	1	THIM_PEA	P48384	pisum sativ
696	5	2.3	159	1	GREB_BUCAI	P57464	buchnera ap	769	5	2.3	173	1	CD3D_RAT	P19377	rattus norv
697	5	2.3	159	1	HS12_DAUCA	P27397	daucus caro	770	5	2.3	173	1	DCD_ACIAM	Q02103	acidianus a
698	5	2.3	159	1	MLE_TODPA	P05945	codarodes p	771	5	2.3	173	1	IPYR_SULAC	P50308	sulfolobus
699	5	2.3	159	1	NUSB_PSEAE	Q52200	pyrus commun	772	5	2.3	173	1	YC37_PORPU	P51191	porphyra su
700	5	2.3	159	1	PYR1_PYRCO	O65200	pyrus commun	773	5	2.3	173	1	YWGJ_BACSU	P06629	bacillus su
701	5	2.3	159	1	XDHG_ECOS7	O846c4	escherichia	774	5	2.3	174	1	ERIL_BOVIN	Q66415	bos taurus
702	5	2.3	159	1	XDHG_ECOLI	O46801	escherichia	775	5	2.3	174	1	MEAI_BOVIN	Q29407	bos taurus
703	5	2.3	160	1	ALCO_PSEAE	P15275	pseudomonas	776	5	2.3	174	1	MSRA_PASMU	Q50871	pasteurella
704	5	2.3	160	1	RS4_SULAC	P39467	sulfolobus	777	5	2.3	174	1	RBS1_WHEAT	Q40004	hordeum vul
705	5	2.3	160	1	SSB_CHILU	O9K24	chlamydia m	778	5	2.3	174	1	RBS_HORVU	P06023	bacterioph
706	5	2.3	160	1	TRIA_TRIPA	O27049	triatoma pa	779	5	2.3	174	1	VGAM_BPMU	O66227	haemophilus
707	5	2.3	160	1	YFOT_ECOLI	P55140	escherichia	780	5	2.3	174	1	Y60A_HAINT	P71901	mycobacteri
708	5	2.3	161	1	DNEI_CHLUV	P56347	chlorella v	781	5	2.3	174	1	YN11_MYCTU	P57115	buchnera ap
709	5	2.3	161	1	IF51_CAEEL	P34563	caenorhabdi	782	5	2.3	175	1	HSLV_BUCAI	P57115	buchnera ap
710	5	2.3	161	1	IF52_CAEEL	Q20751	caenorhabdi	783	5	2.3	175	1	IF3_CHLTR	Q9HNG7	halobacteri
711	5	2.3	161	1	PHAB_AMACY	P07326	anabaena cy	784	5	2.3	175	1	PYRE_HALNT	Q9HNG7	halobacteri
712	5	2.3	161	1	PHAB_AMASP	P80557	anabaena ap	785	5	2.3	175	1	RBS2_WHEAT	P03535	human rofav
713	5	2.3	161	1	PHAB_AMAVA	P00317	anabaena va	786	5	2.3	175	1	IL20_HUMAN	Q9NY11	homo sapien
714	5	2.3	161	1	PHAB_CYACA	O9E188	cyanidium c	787	5	2.3	176	1	IPYR_GLUOX	O05545	gluconobact
715	5	2.3	161	1	PHAB_MASLA	P00318	mastigoclad	788	5	2.3	176	1	IPYR_GLUOX	Q05545	gluconobact
716	5	2.3	161	1	PHAB_SYNEL	P50031	synecococc	789	5	2.3	176	1	NU6C_OENHO	Q6M1C9	neochera h
717	5	2.3	161	1	PHAB_SYNEL	P06113	synecococc	790	5	2.3	176	1	NU6C_TOBAC	Q32722	nicotiana t
718	5	2.3	161	1	PHAB_SYNYA	O01952	synecocyst	791	5	2.3	176	1	RPOE_STRCO	P38123	streptomyce
719	5	2.3	161	1	PHAB_SYNYA	O02824	synecocyst	792	5	2.3	176	1	VATT_ENTDI	Q24808	entamoeba d
720	5	2.3	161	1	PTGA_BUCAI	O9WX17	buchnera ap	793	5	2.3	176	1	V365_MERTU	O57811	methanococc
721	5	2.3	161	1	TPW_SCHPO	Q02088	schizosacch	794	5	2.3	177	1	SHAO_YEAST	P47096	asaccharomyc
722	5	2.3	162	1	C550_CYAPA	P16571	freymyella d	795	5	2.3	177	1	CBP_SACER	P64995	asaccharomyc
723	5	2.3	162	1	PHAB_FREDI	P05730	cyanophora	796	5	2.3	177	1	NUSG_BACSU	Q06793	bacillus su
724	5	2.3	162	1	PHCA_CYAPA	P77484	escherichia	797	5	2.3	177	1	NUSG_PSEAE	O67993	bacillus su
725	5	2.3	162	1	YFHP_ECOLI	P73063	synecocyst	798	5	2.3	177	1	RLIX_SPOER	O8WGI7	spodoptera
726	5	2.3	162	1	Y1S1_SYNY3	P49510	odontella s	799	5	2.3	177	1	VATT_ENTHI	Q24810	entamoeba h
727	5	2.3	163	1	C550_ODOST	P51199	porphyra pu	800	5	2.3	178	1	Y189_STRGR	Q9XHE2	streptomyce
728	5	2.3	163	1	C550_PORPU	O9L1W2	cyanidium c	801	5	2.3	178	1	HPRT_AOUAE	O66821	aquifex aeo
729	5	2.3	164	1	C550_CVACA	P37182	escherichia	802	5	2.3	178	1	IL10_MERUN	P47965	meriones un
730	5	2.3	164	1	HYBD_ECOLI	P36591	onchocerca	803	5	2.3	178	1	IL10_MOUSE	P18893	mus musculu
731	5	2.3	164	1	OV17_ONCVO	O9K9E1	bacillus ha	804	5	2.3	178	1	IL10_RAT	P29456	rattus norv
732	5	2.3	165	1	RL10_BACHD	O9K9E4	bacillus ha	805	5	2.3	178	1	IPYR_AOUAE	O67501	aquifex aeo
733	5	2.3	165	1	RL10_BACSU	P42823	bacillus su	806	5	2.3	178	1	NUSG_BACHD	Q9XGE7	bacillus ha
734	5	2.3	166	1	CYB_DROSU	P51941	drosophila	807	5	2.3	179	1	I122A_MOUSE	Q91JY9	mus musculu
735	5	2.3	166	1	RL14_DROME	P55841	drosophila	808	5	2.3	179	1	I122B_MOUSE	Q91JY8	mus musculu
736	5	2.3	166	1	Y385_ARCTO	O29862	archaeoglob	809	5	2.3	179	1	IL22_HUMAN	O9GZG6	homo sapien
737	5	2.3	166	1	Y504_MCTTU	O11168	mycobacteri	810	5	2.3	179	1	P8SG_RAT	P04917	rattus norv
738	5	2.3	167	1	HRAS_MOUSE	O9GZU4	mus musculu	811	5	2.3	179	1	STC_ONCKE	P33647	oncothychnu
739	5	2.3	168	1	DEP_ECOLI	P27251	escherichia	812	5	2.3	179	1	YC21_ANTSP	Q06463	artichamnio
740	5	2.3	168	1	NEU2_MOUSE	P01486	mus musculu	813	5	2.3	179	1	YCF3_ODOSI	P49525	odontella s
741	5	2.3	168	1	NEU2_RAT	O47478	loiligo blae	814	5	2.3	179	1	YFEL_SALTY	P40196	salmomella
742	5	2.3	168	1	NU6M_LOLBL	O47478	loiligo blae	815	5	2.3	179	1	YHAB_ECOLI	P18865	escherichia
743	5	2.3	168	1	PTGA_ECOLI	P08837	escherichia	816	5	2.3	180	1	CEA5_ECOLI	P18000	escherichia
744	5	2.3	168	1	PTGA_SALTY	P02308	salmomella	817	5	2.3	180	1	FM1C_ECOLI	P08561	escherichia
745	5	2.3	168	1	RL10_RALSO	O8XU26	talstonia s	818	5	2.3	180	1	HBLO_PHPYA	Q06463	physcomitre
746	5	2.3	168	1	YBPO_BACSU	P54158	bacillus su	819	5	2.3	180	1	RBS1_ARATH	P10795	arabidopsis
747	5	2.3	168	1	YU00_AOUAE	O67738	aquifex aeo	820	5	2.3	180	1	SRPA_ECOLI	P12730	escherichia
748	5	2.3	168	1	YMO1_MARPO	P38450	marichantia	821	5	2.3	180	1	YC55_PYRHO	O56998	pyrococcus
749	5	2.3	168	1	YPOE_BACSU	P50829	bacillus su	822	5	2.3	181	1	HSIV_BACSU	P39070	bacillus su
750	5	2.3	169	1	GSHH_ARATH	O48646	arabidopsis	823	5	2.3	181	1	NUSG_ECOLI	P16921	escherichia
751	5	2.3	169	1	GSHH_LYCES	O24031	lycopersico	824	5	2.3	181	1	NUSG_SALTY	Q919X0	salmomella
752	5	2.3	169	1	Y778_METUA	P43608	methanococc	825	5	2.3	181	1	RBS2_ARATH	P10796	arabidopsis
753	5	2.3	169	1	YFJ6_YEAST	O58488	arabidopsis	826	5	2.3	181	1	RBS3_ARATH	P10797	arabidopsis
754	5	2.3	170	1	C13_MATZE	P33050	zea mays (m	827	5	2.3	181	1	RBS4_ARATH	P10798	arabidopsis
755	5	2.3	170	1	GSHH_MESCR	O916F0	mesembryant	828	5	2.3	181	1	RBS_RAPSA	P08135	raphanus sa
756	5	2.3	171	1	AR20_YEAST	P33204	saccharomyc	829	5	2.3	181	1	RNH2_BORBU	O51075	bortrelia bu
757	5	2.3	171	1	BLC_VIBCH	Q08790	vibrio chol	830	5	2.3	181	1	RYVC_RALSO	Q98233	talstonia s
758	5	2.3	171	1	CCAA_CHICK	O73705	gallus gall	831	5	2.3	181	1	TCTP_BRUMA	P09697	brugia mala
759	5	2.3	171	1	PABA_ECOLI	P18391	escherichia	832	5	2.3	181	1	TCTP_WUCA	O96242	whomoceria
760	5	2.3	171	1	PABA_SALTY	O8XGJ3	salmomella	833	5	2.3	181	1	THIM_SPIOL	P05991	spinacia ol
761	5	2.3	171	1	IF3_THRTH	O9AGJ8	thermus the	834	5	2.3	182	1	BLG4_BLABE	P59662	blatella g
762	5	2.3	172	1	PABA_VIBCH	O9K600	vibrio chol	835	5	2.3	182	1	NUSG_STAMM	O08386	staphylococ
763	5	2.3	172	1	PABA_YERPE	O8ZG80	yersinia pe	836	5	2.3	182	1	NUSG_STACA	P36264	staphylococ



837	5	2.3	182	1	NUSG_VIBCH	Q9kx35	vibriol chol	910	5	2.3	197	1	YCB7_PSEDE	P29940	pseudomonas
838	5	2.3	182	1	PYRE_PYRFU	P58861	pyrococcus	911	5	2.3	197	1	YDB6_YEAST	Q12055	saccharomyc
839	5	2.3	182	1	R104_YEAST	P33321	saccharomyc	912	5	2.3	198	1	GRPE_HABIN	P33732	haemophilus
840	5	2.3	182	1	VG45_HAEIN	P44282	haemophilus	913	5	2.3	198	1	LEUD_MYCTE	Q31314	mycobacteri
841	5	2.3	182	1	YGF8_HAEIN	P44882	haemophilus	914	5	2.3	198	1	LEUD_MYCTE	Q53236	mycobacteri
842	5	2.3	182	1	YN83_SUTTO	Q96xy5	sulfobius	915	5	2.3	198	1	RECR_CLOAB	Q97m44	clostridium
843	5	2.3	182	1	YON6_CABEL	Q09527	caenorhabdi	916	5	2.3	198	1	RECR_CLOAB	Q22769	listeria in
844	5	2.3	183	1	ATKC_PSEAE	P57686	pseudomonas	917	5	2.3	198	1	RECR_LISIN	Q93j47	listeria in
845	5	2.3	183	1	CBX3_HUMAN	Q13186	homo sapien	918	5	2.3	198	1	RNM2_ECOLI	P10442	escherichia
846	5	2.3	183	1	CBX3_MOUSE	P23198	mus musculus	919	5	2.3	198	1	RNM2_ECOLI	P40652	escherichia
847	5	2.3	183	1	NUSG_PASMU	Q9ck84	pasteurella	920	5	2.3	198	1	VS11_PORRA	P17467	salmonella
848	5	2.3	184	1	LBP_RENRE	P05938	renilla ren	921	5	2.3	198	1	Y530_MERUA	P17950	rabbit rota
849	5	2.3	184	1	NUSG_BORBU	Q51355	borrelia bu	922	5	2.3	199	1	CAGS_HELPJ	Q57960	methanococ
850	5	2.3	184	1	NUSG_THETH	P35872	thermus the	923	5	2.3	199	1	HIRS_MOUSE	Q9zj14	helicobacte
851	5	2.3	185	1	NO20_SOYBN	P08960	glycine max	924	5	2.3	199	1	HM19_CABEL	Q6q273	mus musculu
852	5	2.3	185	1	NUSG_HAEIN	P43916	haemophilus	925	5	2.3	199	1	RABA_DICDI	P44141	dictyosteli
853	5	2.3	185	1	RRF_STRPY	P82556	streptococc	926	5	2.3	199	1	RSBX_BACSU	P17966	bacillus su
854	5	2.3	185	1	US10_HCMVA	P09728	human cytom	927	5	2.3	199	1	SODF_RHIME	Q9xkd4	rhizobium m
855	5	2.3	185	1	Y920_MERUA	Q58330	methanococc	928	5	2.3	199	1	WRB1_RHIME	Q32p33	rhizobium m
856	5	2.3	185	1	YE46_AOUAE	Q67433	aquifex aeo	929	5	2.3	199	1	YBT9_YEAST	P38252	saccharomyc
857	5	2.3	185	1	YRKN_BACSU	P54441	bacillus su	930	5	2.3	199	1	YET4_HABIN	Q57213	haemophilus
858	5	2.3	186	1	ATPD_RHOBL	P05437	rhodospseudo	931	5	2.3	199	1	YPHA_BACSU	P50741	bacillus su
859	5	2.3	187	1	OM24_ARATH	P82805	arabidopsis	932	5	2.3	200	1	COAE_MYCPN	P75400	mycoplasma
860	5	2.3	187	1	PNOC_MOUSE	Q64387	mus musculu	933	5	2.3	200	1	R102_YEAST	Q02721	saccharomyc
861	5	2.3	187	1	Y893_HAEIN	P44923	haemophilus	934	5	2.3	200	1	STRM_STRGR	P29763	streptomyce
862	5	2.3	187	1	YM95_MYCTU	Q50673	mycobacteri	935	5	2.3	201	1	RB9B_HUMAN	Q09p30	homo sapien
863	5	2.3	188	1	DCD_PSEAE	Q9hy93	pseudomonas	936	5	2.3	201	1	RNM2_MAGSA	Q50412	magnetospir
864	5	2.3	188	1	R1MM_AGR75	Q8ubz8	agrobacteri	937	5	2.3	201	1	RUVV_LISMO	Q8Y657	listeria mo
865	5	2.3	188	1	YVDE_LACTC	P22347	lactococcus	938	5	2.3	202	1	R13A_CABEL	Q27369	caenorhabdi
866	5	2.3	189	1	CHS1_TUBUN	P55003	tuber uncin	939	5	2.3	202	1	RPOW_SULAC	Q07271	bulboblus
867	5	2.3	189	1	PTGA_BORBU	Q44840	borrelia bu	940	5	2.3	202	1	RS2_CRIGR	Q07271	bulboblus
868	5	2.3	189	1	TBP_THECE	Q56253	thermococcu	941	5	2.3	202	1	UMPK_ARATH	Q04905	arabidopsis
869	5	2.3	190	1	NUSB_MYCLC	Q9ccr9	mycobacteri	942	5	2.3	202	1	VNS1_IAROM	P53459	influenza a
870	5	2.3	190	1	RNM2_SYNY3	P72657	synthocycst	943	5	2.3	202	1	Y038_MYCLE	Q50191	mycobacteri
871	5	2.3	190	1	RS7_AVIMR	Q9zne1	avicennia m	944	5	2.3	202	1	Y1BF_ECOLI	P32105	escherichia
872	5	2.3	190	1	SLVD_HABIN	P44830	haemophilus	945	5	2.3	203	1	BET1_RHIME	Q69786	rhizobium m
873	5	2.3	190	1	TBP_PYROO	Q52366	pyrococcus	946	5	2.3	203	1	FIBR_AGRKO	P28891	agkistrodon
874	5	2.3	190	1	Y052_HAANI	Q9hmp0	halobacteri	947	5	2.3	203	1	FMS5_ECOLI	P33761	escherichia
875	5	2.3	191	1	CBX5_HUMAN	P45973	homo sapien	948	5	2.3	203	1	GTS2_MANSE	P63479	manduca sex
876	5	2.3	191	1	CBX5_MOUSE	O61666	mus musculu	949	5	2.3	203	1	RAB7_DICDI	P64411	dictyosteli
877	5	2.3	191	1	NUSC_MARPO	P06266	marichantia	950	5	2.3	203	1	VP10_BPPOD	P28712	bacterioph
878	5	2.3	191	1	TBP_PYRAB	Q9V024	pyrococcus	951	5	2.3	203	1	Y461_PYRHO	P58824	pyrococcus
879	5	2.3	191	1	TBP_PYRHO	O58737	pyrococcus	952	5	2.3	203	1	YJ86_ARCFU	Q28223	archaeoglob
880	5	2.3	191	1	TNR6_ENTFC	Q06237	enterococcu	953	5	2.3	204	1	CRB2_BOVIN	P25252	bos taurus
881	5	2.3	191	1	YJDC_ECOLI	P36656	escherichia	954	5	2.3	204	1	CRB2_HUMAN	P33320	homo sapien
882	5	2.3	192	1	G16U_BACSU	P80875	bacillus su	955	5	2.3	204	1	CRB2_MOUSE	P62755	mus musculu
883	5	2.3	192	1	RAS2_HYDMA	P38976	hydra magni	956	5	2.3	204	1	CYPS_CABEL	P52013	caenorhabdi
884	5	2.3	192	1	WMT2_LABMD	P03721	bacterioph	957	5	2.3	204	1	EVGA_ECOLI	P30854	schizophy11
885	5	2.3	193	1	LEP2_BACAM	P41025	bacillus am	958	5	2.3	204	1	SC7_SCHCO	P35794	schizophy11
886	5	2.3	193	1	RNM2_RICCN	Q92105	rickettsia	959	5	2.3	204	1	UREG_SFRSL	O55057	streptococc
887	5	2.3	193	1	YLB0_BACSU	O34549	bacillus su	960	5	2.3	204	1	YVPE_BACSU	P39638	bacillus su
888	5	2.3	194	1	FGP7_SHEEP	P48808	ovis aries	961	5	2.3	205	1	ALKH_ZYMOO	Q00394	z kbg/kdpg
889	5	2.3	194	1	IL18_RAT	P97636	rattus norv	962	5	2.3	205	1	CD83_HUMAN	Q01151	homo sapien
890	5	2.3	194	1	KAD_BRUME	Q8yh19	bruceella me	963	5	2.3	205	1	MTR3_HUMAN	Q13615	homo sapien
891	5	2.3	194	1	RNM2_NEIMA	Q9jx40	neisseria m	964	5	2.3	205	1	RACI_DICDI	Q9p922	dictyosteli
892	5	2.3	194	1	RNM2_NEIMB	Q9k1g1	neisseria m	965	5	2.3	205	1	RNM2_CAUCR	P32975	caulobacter
893	5	2.3	194	1	UREF_HAEIN	P44395	haemophilus	966	5	2.3	205	1	YPT7_SCHPO	O94655	schistosach
894	5	2.3	195	1	AAC2_MYCRO	Q49157	mycobacteri	967	5	2.3	206	1	KAD_AOUAE	O66400	aquifex aeo
895	5	2.3	195	1	KTHY_AOUAE	O67099	aquifex aeo	968	5	2.3	206	1	NHAA_RHOER	P13448	rhodococcus
896	5	2.3	195	1	VAIE_HAANI	Q9hne0	halobacteri	969	5	2.3	206	1	RALA_HUMAN	P11233	homo sapien
897	5	2.3	195	1	YL47_ARCFU	O28135	archaeoglob	970	5	2.3	206	1	RALA_MOUSE	P05810	mus musculu
898	5	2.3	196	1	CAGS_HELPJ	P97227	helicobacte	971	5	2.3	206	1	RALB_HUMAN	P11224	homo sapien
899	5	2.3	196	1	HIRS_HUMAN	Q9um20	homo sapien	972	5	2.3	206	1	RALB_RAT	P36860	rattus norv
900	5	2.3	196	1	HIS7_RALSO	O8kx01	raistrontia s	973	5	2.3	206	1	RAL_DICOM	P22124	discozyme o
901	5	2.3	196	1	RNM2_AOUAE	O67768	aquifex aeo	974	5	2.3	206	1	RNFG_SALTI	O8z677	salmonella
902	5	2.3	196	1	SPMA_BACSU	P51517	bacillus su	975	5	2.3	206	1	RNM2_VIBCH	P52021	vibrio chol
903	5	2.3	196	1	Y546_NEIMA	Q9jw50	neisseria m	976	5	2.3	206	1	RS4_PSEAE	O52759	pseudomonas
904	5	2.3	196	1	YJ09_NEIMB	O9jxw2	neisseria m	977	5	2.3	207	1	ACPD_URBPA	O56551	ureaplasma
905	5	2.3	197	1	ATP4_PEA	Q41000	pisum sativ	978	5	2.3	207	1	HISS_AOUAE	O8Yv96	anabena sp
906	5	2.3	197	1	IP21_TOBAC	Q40561	nicotiana t	979	5	2.3	207	1	LOLB_ECOLI	O66993	aquifex aeo
907	5	2.3	197	1	MAUD_PARVE	Q56461	paracoccus	980	5	2.3	207	1	LOLB_ECOLI	P24208	escherichia
908	5	2.3	197	1	RNM2_HAEIN	P43808	haemophilus	981	5	2.3	207	1	LOLB_SALTY	P30752	salmonella
909	5	2.3	197	1	RNM2_PASMU	P57966	pasteurella	982	5	2.3	207	1	PSB3_GIALA	Q9ngw8	giardia lam



```

983 5 2.3 207 1 RAB7 CANPA
984 5 2.3 207 1 RAB7 HUMAN
985 5 2.3 207 1 RAB7 MOUSE
986 5 2.3 207 1 RAB7_RABIT
987 5 2.3 207 1 RAB7_RAT
988 5 2.3 207 1 RAB7_CHLVU
989 5 2.3 207 1 VF01 VACCV
990 5 2.3 207 1 Y078 METJA
991 5 2.3 207 1 YP90 PASMU
992 5 2.3 208 1 CLDY_BRARE
993 5 2.3 208 1 CSF3_MOUSE
994 5 2.3 208 1 ENGB_UREPA
995 5 2.3 208 1 PYRF_PYRAB
996 5 2.3 208 1 RS4_HELPJ
997 5 2.3 208 1 RS4_HELPJ
998 5 2.3 208 1 YS22_CAEEL
999 5 2.3 209 1 BEL6_MOUSE
1000 5 2.3 209 1 CLDZ_BRARE

```

## ALIGNMENTS

```

RESULT 1
OSCI BORBU STANDARD: PRT: 210 AA.
ID OSCI BORBU
AC 007337:
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Outer surface protein C precursor (PC).
GN OSPC OR BBH19.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Plasmid lps4.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCB1_Taxid=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=93268136; PubMed=8098841;
RA Jauris-Heipke S., Fuchs R., Motz M., Preac-Mursic V., Schwab E.,
RA Will G., Wilske B.;
RT "Genetic heterogeneity of the genes coding for the outer surface
RT protein C (OSPC) and the flagellin of Borrelia burgdorferi.";
RL Med. Microbiol. Immunol. 182:37-50(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=93268136; PubMed=8478108;
RA Wilske B., Preac-Mursic V., Jauris S., Pradel I., Soutschek E.,
RA Schwab E., Manner G.;
RT "Immunological and molecular polymorphisms of OspC, an immunodominant
RT major outer surface protein of Borrelia burgdorferi.";
RL Infect. Immun. 61:2182-2191(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=94041630; PubMed=8225587;
RA Padula S.J., Sampliner A., Dias F., Szczepanski A., Ryan R.W.;
RT "Molecular characterization and expression of p23 (OspC) from a North
RT American strain of Borrelia burgdorferi.";
RL Infect. Immun. 61:5097-5105(1993).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=96025162; PubMed=7494039;
RA Fukunaga M., Hamase A.;
RT "Outer surface protein C gene sequence analysis of Borrelia
RT burgdorferi sensu lato isolates from Japan.";
RL J Clin. Microbiol. 33:2415-2420(1995).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;

```

```

RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lachlani R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kertlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt J., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Uiterback T., Wiethe L., McDonald L., Artach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia
RT burgdorferi";
RL Nature 390:580-586(1997).
CC -1- FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.
CC -1- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC anchor.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; X69596; CAA49306.1; -
DR EMBL; U01894; AAA16058.1; -
DR EMBL; D49497; BAA08457.1; -
DR EMBL; AE000792; AAC6329.1; -
DR TIGR; BBH19; -
DR InterPro; IPR001800; Lipoprotein_6.
DR Pfam; PF01441; Lipoprotein_6; 1.
DR ProDom; PD001149; Lipoprotein_6; 1.
DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.
KW Outer membrane; Lipoprotein; Signal; Plasmid; Antigen;
KW Complete proteome.
KM BY SIMILARITY.
FT SIGNAL 1 18
FT CHAIN 19 210 OUTER SURFACE PROTEIN C.
FT LIPID 19 N-ACYL DIGLYCERIDE (BY SIMILARITY).
SQ SEQUENCE 210 AA; 22340 MW; 7A4FC978F91777BF CRC64;

Query Match 3.2%; Score 7; DB 1; Length 210;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 39 AKELIK 45
Db 155 AKELIK 161

RESULT 2
HIS1_PSEAE
ID HIS1_PSEAE STANDARD: PRT: 213 AA.
AC 09HU42;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Imidazole glycerol phosphate synthase subunit hisH1 (EC 2.4.2.-) (IGP
DE synthase glutamine amidotransferase subunit) (IGP synthase subunit
DE hisH1) (ImGP synthase subunit hisH1) (IGPS subunit hisH1).
GN HIS1 OR PA5142.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCB1_Taxid=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Britman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,
RA Garner R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lardys K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,

```

```
RA Reizer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;  
RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an  
opportunistic pathogen."  
RL Nature 406:959-964(2000).  
CC  
CC -1- FUNCTION: IGP5 catalyzes the conversion of PRPP and glutamine to  
ICP, AICAR and glutamate. The high subunit provides the glutamine  
and/or transaminase activity that produces the ammonia necessary to  
hisp for the synthesis of IGP and AICAR (By similarity).  
CC  
CC -1- CATALYTIC ACTIVITY: 5-(1-(5-phospho-1-deoxyribulos-1-  
ylamino)methylideneamino)-1-(5-phosphoribosyl)imidazole-4-  
carboxamide + L-glutamine = imidazole-glyceroi phosphate + 5-  
antioimidazole-4-carboxamide ribonucleotide + L-glutamate + H(2)O.  
CC  
CC -1- PATHWAY: Histidine biosynthesis; fifth step.  
CC  
CC -1- SUBUNIT: Heterodimer of hisH and hisF (By similarity).  
CC  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC  
CC -1- SIMILARITY: CONTAINS 1 TYPE-1 GLUTAMINE AMIDOTRANSFERASE DOMAIN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC  
CC -----  
CC EMBL; AE004927; AAC08527.1; -.  
CC InterPro: IPR000991; GATase_1.  
CC Pfam; PF00117; GATase_1.  
CC PROSITE; PS00442; GATASE_TYPE_1; 1.  
CC Histidine biosynthesis; Transferrase; Glutamine amidotransferase;  
CC Complete proteome.  
CC  
CC ACT_SITE 82 BY SIMILARITY.  
CC FT ACT_SITE 191 BY SIMILARITY.  
CC FT ACT_SITE 193 BY SIMILARITY.  
CC SQ SEQUENCE 213 AA; 23698 MW; ECE247CDB0411F8 CRC64;  
  
Query Match 3.2%; Score 7; DB 1; Length 213;  
Best local Similarity 100.0%; Pred. No. 11;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 179 GLOLON 185  
Db 199 GLOLON 205  
  
RESULT 3  
ABC_ECOLI STANDARD; PRT; 343 AA.  
AC P30750; P77517;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE ATP-binding protein abc.  
GN abc OR B0199.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12 / MG1655;  
RX MEDLINE=9742617; PubMed=9278503;  
RA Blattner F.R., Plunkett G. II, Bloch C.A., Perna N.T., Burland V.,  
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
RA Ma B., Siao Y.;  
RT "The complete genome sequence of Escherichia coli K-12."  
RL Science 277:1453-1474(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12 / W3110;  
RA Takemoto K., Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,  
RA Yamamoto Y., Inokuchi H., Miki T., Hatada E., Fukuda R., Ichihara S.,
```

```
RA Mizuno T., Makino K., Nakata A., Yura T., Sampei G., Mizobuchi K.;  
RT "Systematic sequencing of the Escherichia coli genome: analysis of the  
4.0 - 6.0 min (189,987 - 281,416bp) region."  
RL Submitted (FE8-1996) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E.,  
RA Davis K., Federpiel N., Hyman R., Kalman S., Komp C., Kurd O.,  
RA Laebhar D., Lew H., Lin D., Namath A., Oefner P., Roberts D.,  
RA Davis R.W.;  
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE OF 1-231 FROM N.A.  
RC STRAIN=TAP90;  
RX MEDLINE=9412404; PubMed=7904973;  
RA Allikmets R., Gerrard B.C., Court D., Dean M.C.;  
RT "Cloning and organization of the abc and mdl genes of Escherichia  
coli: relationship to eukaryotic multidrug resistance."  
RL Gene 136:231-236(1993).  
CC  
CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).  
CC  
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.  
CC  
CC -1- CAUTION: REF.4 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A  
CC FRAMESHIFT IN POSITION 200.  
CC  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC  
CC -----  
CC EMBL; AE000129; AAC73310.1; -.  
CC EMBL; D83536; BAA77876.1; -.  
CC DR EMBL; U70214; AAB08627.1; -.  
CC DR EMBL; L08626; AAC36869.1; ALT_FRAME.  
CC DR HSP; Q58663; IG6H.  
CC DR Ecocore; EGI1621; abc.  
CC DR InterPro: IPR003593; AAA_ATPase.  
CC DR InterPro: IPR003439; ABC_transporter.  
CC Pfam; PF00005; ABC_tran; 1.  
CC ProDom; PD000006; ABC_transporter; 1.  
CC SMART; SM00382; AAA; 1.  
CC PROSITE; PS00211; ABC_TRANSPORTER; 1.  
CC KM Transport; ATP-binding; Inner membrane; Complete proteome.  
CC NP BIND 38 45 ATP (POTENTIAL).  
CC FT CONFLICT 106 106 P -> R (IN REF. 2).  
CC FT CONFLICT 165 165 D -> V (IN REF. 4).  
CC SQ SEQUENCE 343 AA; 37788 MW; 2FC796C605216FB3 CRC64;  
  
Query Match 3.2%; Score 7; DB 1; Length 343;  
Best local Similarity 100.0%; Pred. No. 18;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 173 GSVLVDG 179  
Db 60 GSVLVDG 66  
  
RESULT 4  
Y4TL_RHISN STANDARD; PRT; 390 AA.  
ID Y4TL_RHISN  
AC P55666;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Hypothetical hydrolase/peptidase Y4TL (EC 3.-.-.-).  
GN Y4TL.  
OS Rhizobium sp. (strain NGR234).  
OC Plasmid sym pNGR234a.  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Rhizobium.  
OX NCBI_TaxID=394;
```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97305556; PubMed=9163424;
RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
RA Perret X.;
RT "Molecular basis of symbiosis between Rhizobium and legumes.";
RL Nature 387:394-401(1997).
CC -1- SIMILARITY: BELONGS TO THE SUPERFAMILY OF PEPTIDASES M24 WHICH
CC GROUPS MEMBRANE AMINOPEPTIDASES (M24A/M24C), XAA-PRO
CC AMINOPEPTIDASES (M24B) AND CREATININASES.
CC -1- SIMILARITY: TO Y4TM.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000098; AAB91865.1; -
DR MEROPS; M24.UNB; -
DR InterPro; IPR000994; Peptidase_M24.
DR Pfam; PF00557; Peptidase_M24; 1.
DR Hypothetical protein; Hydrolase; Plasmid.
SQ SEQUENCE 390 AA; 42773 MW; 9213FBC2E0AFE4D CRC64;

Query Match 3.2%; Score 7; DB 1; Length 390;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 86 IKSDLEI 92
Db 162 IKSDLEI 168

RESULT 5
PGK_THETN STANDARD; PRT; 394 AA.
AC Q8R65;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phosphoglycerate kinase (EC 2.7.2.3).
GN PGK OR TTB1761.
OS Thermomicrobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermomicrobacteriales;
OC Thermomicrobacteriaceae; Thermomicrobacter.
OX NCBI_Taxid=119072;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=MBA / JCM 11007;
RX MEDLINE=21992816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of T. tengcongensis genome.";
RL Genome Res. 12:689-700(2002).
CC -1- CATALYTIC ACTIVITY: ATP + 3-phospho-D-glycerate = ADP + 3-
CC phospho-D-glycerol phosphate.
CC -1- PATHWAY: Second phase of glycolysis; second step.
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE PHOSPHOGLYCERATE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

```

DR EMBL; AE013129; AAM24955.1; -
DR PROSITE; PS00111; GLYCERATE KINASE; 1.
KW Transferase; Kinase; Glycolysis; Complete proteome.
SQ SEQUENCE 394 AA; 42816 MW; 983EA9E59C240A32 CRC64;

Query Match 3.2%; Score 7; DB 1; Length 394;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 35 KIELAKE 41
Db 243 KIELAKE 249

RESULT 6
PGK_SYNY3 STANDARD; PRT; 401 AA.
AC P74421;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phosphoglycerate kinase (EC 2.7.2.3).
GN PGK OR SLR0394.
OS Synecocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
OX NCBI_Taxid=1148;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugita M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Nario K.,
RA Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,
RA Yamada M., Yasuda M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synecocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
CC -1- CATALYTIC ACTIVITY: ATP + 3-phospho-D-glycerate = ADP + 3-
CC phospho-D-glycerol phosphate.
CC -1- PATHWAY: Second phase of glycolysis; second step.
CC -1- SUBUNIT: MONOMER (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE PHOSPHOGLYCERATE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D90915; BAA18521.1; ALT_INIT.
DR HSP; P18912; IPRP.
DR InterPro; IPR001576; PGK.
DR Pfam; PF00162; PGK; 1.
DR PRINTS; PR00477; PHGLYCKINASE.
DR PROSITE; PS00111; GLYCERATE KINASE; 1.
KW Transferase; Kinase; Glycolysis; Complete proteome.
SQ SEQUENCE 401 AA; 41784 MW; 74905FDBCBF1B7D CRC64;

Query Match 3.2%; Score 7; DB 1; Length 401;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 95 RUTPVGD 101
Db 76 RUTPVGD 82

RESULT 7
SYN_METTH
```

```

ID  SVP_METHH  STANDARD;  PRT;  482 AA.
AC  026708;
DT  15-JUL-1998 (Rel. 36, Created)
DT  15-JUL-1998 (Rel. 36, Last sequence update)
DT  15-OCT-2001 (Rel. 40, Last annotation update)
DE  Prolyl-tRNA synthetase (EC 6.1.1.15) (Proline--tRNA ligase) (Profs).
GN  PROS OR MTH611.
OS  Methanobacterium thermoautotrophicum.
OC  Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC  Methanobacteriaceae; Methanothermobacter.
OX  NCBI_TaxID=187420;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=Delta H;
RX  MEDLINE=98037514; PubMed=9371463;
RA  Smith D.R., Doucet-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA  Aldredge T., Bahrzad R., Blakely D., Cook R., Gilbert K.,
RA  Harrison D., Hoang L., Keagle P., Lumm M., Pochler B., Qiu D.,
RA  Spadator R., Viccare R., Wang Y., Wierzbowski J., Gibson R.,
RA  Jiwani N., Carnuo A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA  McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA  Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.,
RT  "Complete genome sequence of Methanobacterium thermoautotrophicum
RT  deltaH: functional analysis and comparative genomics.";
RL  J. Bacteriol. 179:7135-7155(1997).
CC  -1- CATALYTIC ACTIVITY: ATP + L-proline + tRNA(Pro) = AMP +
CC  diphosphate + L-prolyl-tRNA(Pro).
CC  -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC  -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AE000842; AAB85117.1; -.
DR  InterPro; IPR002106; AATRNA_ligase1.
DR  InterPro; IPR004154; HGTP_anticon.
DR  InterPro; IPR004499; ProS_fam_1.
DR  InterPro; IPR002316; tRNA-synt_2b.
DR  InterPro; IPR002316; tRNA-synt_pro.
DR  Pfam; PF00587; tRNA-synt_2b; 1.
DR  Pfam; PF01129; HGTP_anticon; 1.
DR  PRINTS; PR01046; TRNASYNTHPRO.
DR  TIGRFAMs; TIGR00408; ProS_fam_1; 1.
DR  PROSITE; PS50862; AA-TRNA_LIGASE_II; 1.
KM  Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KM  Complete proteome.
SQ  SEQUENCE 482 AA; 55805 MW; 57D81ED7B9496BAA CRC64;

Query Match 3.2%; Score 7; DB 1; Length 482;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 ELAKEA1 43
DB 82 ELAKEA1 88

```

```

OX  NCBI_TaxID=10904;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=89073916; PubMed=2849255;
RA  Gould A.R., Pritchard L.I., Tavaris M.D.;
RT  "Nucleotide and deduced amino acid sequences of the non-structural
RT  protein, NS1, of Australian and South African bluetongue virus
RT  serotype 1.";
RL  Virus Res. 11:97-107(1988).
DR  PIR; A60000; A60000.
DR  InterPro; IPR002630; Orbl_NSI.
DR  Pfam; PF01718; Orbl_NSI; 1.
KM  Nonstructural protein.
SQ  SEQUENCE 552 AA; 64593 MW; D96BFA25587463D1 CRC64;

Query Match 3.2%; Score 7; DB 1; Length 552;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 IELAKEA 42
DB 68 IELAKEA 74

RESULT 9
VNS1_BT20  STANDARD;  PRT;  552 AA.
AC  P32931;
DT  01-OCT-1993 (Rel. 27, Created)
DT  01-OCT-1993 (Rel. 27, Last sequence update)
DT  15-DEC-1998 (Rel. 37, Last annotation update)
DE  Nonstructural protein NS1.
GN  S5 OR S6.
OS  Bluetongue virus (serotype 20 / isolate Australia).
OC  Viruses; dsRNA viruses; Reoviridae; Orbivirus.
OX  NCBI_TaxID=31562;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=93118247; PubMed=1335630;
RA  Cowley J.A.;
RT  "Nucleotide sequence of the genome segment encoding nonstructural
RT  protein NS1 of bluetongue virus serotype 20 from Australia.";
RL  Virus Genes 6:387-392(1992).
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; X56735; CAA40059.1; -.
DR  PIR; S25105; S25105.
DR  PIR; A48553; A48553.
DR  InterPro; IPR002630; Orbl_NSI.
DR  Pfam; PF01718; Orbl_NSI; 1.
KM  Nonstructural protein.
SQ  SEQUENCE 552 AA; 64511 MW; 9B72CF6B1F9F4993 CRC64;

Query Match 3.2%; Score 7; DB 1; Length 552;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 IELAKEA 42
DB 68 IELAKEA 74

RESULT 10
G6PI_YEAST  STANDARD;  PRT;  553 AA.
ID  G6PI_YEAST
AC  P12709;

```

DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Glucose-6-phosphate isomerase (EC 5.3.1.9) (GPI) (Phosphoglucose  
 DE isomerase) (PFI) (Phosphohexose isomerase) (PHI).  
 GN PGI1 OR YBR196C OR YBR1406.  
 OS *Saccharomyces cerevisiae* (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 NCBI\_TaxID=4932;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=89211945; PubMed=3072254;  
 RA Tekamp-Olson P., Najarian R., Burke R.L.;  
 RT "The isolation, characterization and nucleotide sequence of the  
 RT phosphoglucose isomerase gene of *Saccharomyces cerevisiae*.";  
 RL Gene 73:153-161(1988).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=AB320;  
 RX MEDLINE=89201230; PubMed=3071735;  
 RA Green J.B.A., Wright A.P.H., Cheung W.Y., Lancashire W.E.,  
 RA Hartley B.S.;  
 RT "The structure and regulation of phosphoglucose isomerase in  
 RT *Saccharomyces cerevisiae*.";  
 RL Mol. Gen. Genet. 215:100-106(1988).  
 [3]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=S288C;  
 RX MEDLINE=93348777; PubMed=8346681;  
 RA Demolis N., Mallet L., Bussereau F., Jacquet M.;  
 RT "RM2, MS1 and PGI1 are located within an 8 kb segment of  
 RT *Saccharomyces cerevisiae* chromosome II, which also contains the  
 RT putative ribosomal gene I21 and a new putative essential gene with a  
 RT leucine zipper motif.";  
 RL Yeast 9:645-659(1993).  
 [4]  
 RN SEQUENCE OF 423-553 FROM N.A.  
 RP STRAIN=S288C;  
 RX MEDLINE=9506384; PubMed=7975899;  
 RA Mallet L., Bussereau F., Jacquet M.;  
 RT "Nucleotide sequence analysis of an 11.7 kb fragment of yeast  
 RT chromosome II including BEM1, a new gene of the WD-40 repeat family  
 RT and a new member of the KRE2/MNT1 family.";  
 RL Yeast 10:819-831(1994).  
 [5]  
 RN ACETYLATION.  
 RA Garrels J.I., Fitcher B., Kobayashi R., Latter G.I., Schwenker B.,  
 RA Volpe T., Warner J.R., McLaughlin C.S.;  
 RL Submitted (SEP-1994) to the SWISS-PROT data bank.  
 CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-  
 CC phosphate.  
 CC -1- PATHWAY: Involved in glycolysis and in gluconeogenesis.  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO THE GPI FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M37267; AAA34862.1; -  
 DR EMBL; M21696; AAA34894.1; -  
 DR EMBL; X13977; CAA32158.1; -  
 DR EMBL; Z21487; CAA79683.1; -  
 DR EMBL; Z36065; CAA85158.1; -  
 DR PIR; J70484; NUBY.  
 DR HSSP; O9N1E2; IHGX.  
 DR SGD; S0000400; PGI1.

DR InterPro: IPR001672; G6P\_Isomerase.  
 DR Pfam: PF00342; PGI.1.  
 DR PRINTS: PR00662; G6PISOMERASE.  
 DR PROSITE: PS00765; P\_GLUCOSE\_ISOMERASE\_1; 1.  
 DR PROSITE: PS00174; P\_GLUCOSE\_ISOMERASE\_2; 1.  
 KM Gluconeogenesis; Glycolysis; Isomerase; Acetylation.  
 FT INT MET 0 0  
 FT ACT SITE 397 397 BY SIMILARITY.  
 FT ACT SITE 519 519 BY SIMILARITY.  
 FT MOD\_RES 1 1 ACETYLATION.  
 SQ SEQUENCE 553 AA; 61168 MW; 55775F78452E1952 CRC64;  
 Query Match 3.2%; Score 7; DB 1; Length 553;  
 Best local Similarity 100.0%; Pred. No. 28;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 36 IELAKEA 42  
 Db 77 IELAKEA 83  
 RESULT 11  
 YBUL\_ECOLI STANDARD; PRT; 561 AA.  
 AC P75812; P71206;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hypothetical membrane protein ybJL.  
 GN YBUL OR B0847 OR Z1074 OR EC50927.  
 OS *Escherichia coli*, and  
 OS *Escherichia coli* O157:H7.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 NCBI\_TaxID=562, 83334;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-VIDES J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of *Escherichia coli* K-12.";  
 RL Science 277:1453-1474(1997).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=K12;  
 RX MEDLINE=97061202; PubMed=8905232;  
 RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,  
 RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,  
 RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,  
 RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,  
 RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,  
 RA Yano M., Horiuchi T.;  
 RT "A 718-kb DNA sequence of the *Escherichia coli* K-12 genome  
 RT corresponding to the 12.7-28.0 min region on the linkage map.";  
 RL DNA Res. 3:137-155(1996).  
 [3]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=O157:H7 / EDL933 / ATCC 700927;  
 RC MEDLINE=21074935; PubMed=11206551;  
 RX Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
 RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
 RA Grodzick E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,  
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
 RA Welch R.A., Blattner F.R.;  
 RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.";  
 RL Nature 409:529-533(2001).  
 [4]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=O157:H7 / RIMD 0509952;

```

RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasekawa C., Ogasawara N., Yasunaga T.,
RA Kohara S., Shiba T., Hattori M., Shinagawa H.;
RT "complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22(2001).
RN
RN SEQUENCE OF 396-527 FROM N.A.
RC STRAIN=K12 / AB1157;
RX MEDLINE=96345615; PubMed=9755878;
RA Zenno S., Koike H., Kumar A.N., Jayaraman R., Tanokura M., Saigo K.;
RT Biochemical characterization of NfsA, the Escherichia coli major
RT nitroreductase exhibiting a high amino acid sequence homology to Ftp,
RT a Vibrio Harvey flavin oxidoreductase."
RL J. Bacteriol. 178:4508-4514(1996).
CC
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE DPF0199 FAMILY. YBL SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AE000187; AAC73934.1; -.
CC EMBL; D90722; BAA35551.1; -.
CC EMBL; D90723; BAA35559.1; -.
CC EMBL; AE005266; AAG55223.1; -.
CC EMBL; AP002553; BAB34350.1; -.
CC EMBL; D38308; BAA07424.1; -.
CC DR Ecogen; EG31681; ybJL.
CC DR InterPro: IPR000308; TrkA_KupTake.
CC DR Pfam; PF02080; TrkA-C; 2.
CC KW Hypothetical protein; Transmembrane; Complete proteome.
CC ID ESTU MANSE 13 35 POTENTIAL.
CC FT TRANSMEM 64 86 POTENTIAL.
CC FT TRANSMEM 93 115 POTENTIAL.
CC FT TRANSMEM 159 181 POTENTIAL.
CC FT TRANSMEM 383 402 POTENTIAL.
CC FT TRANSMEM 406 428 POTENTIAL.
CC FT TRANSMEM 449 471 POTENTIAL.
CC FT TRANSMEM 476 498 POTENTIAL.
CC FT TRANSMEM 535 557 POTENTIAL.
CC FT CONFLICT 522 526 ISDTA -> TLKCI (IN REF. 5).
CC SQ SEQUENCE 561 AA; 60351 MW; 9FD6F8CA67945F85 CRC64;

Query Match 3.2%; Score 7; DB 1; Length 561;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 204 VTEERVV 210
Db 303 VTEERVV 309

RESULT 12
YABU_RHISN
ID YABU_RHISN STANDARD; PRT; 630 AA.
AC P55377;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical 67.9 kDa protein YABU.
GN YABU.
OS Rhizobium sp. (strain NGR234).
OC Plasmid sym pNGR234a.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OC NCBI_TaxID=394;

```

```

RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=97305956; PubMed=9163424;
RA Perleberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
RA Parry X.;
RT "Molecular basis of symbiosis between Rhizobium and legumes."
RL Nature 387:394-401(1997).
CC
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: DISTANTLY RELATED TO PEPTIDASE FAMILY S2C.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AE000066; AAB91625.1; -.
CC DR MEROPS; S01.UPC; -.
CC DR InterPro: IPR001254; Ser protease_Try.
CC Pfam; PF00089; trypsin; 1.
CC KW Hypothetical protein; Plasmid; Transmembrane.
CC FT TRANSMEM 8 28 POTENTIAL.
CC FT TRANSMEM 258 278 POTENTIAL.
CC SQ SEQUENCE 630 AA; 67854 MW; 4C1FB7BA2EB44A02 CRC64;

Query Match 3.2%; Score 7; DB 1; Length 630;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 LYTFAPY 65
Db 263 LYTFAPY 269

RESULT 13
ESTU MANSE
ID ESTU MANSE STANDARD; PRT; 15 AA.
AC P19985;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DE 01-FEB-1994 (Rel. 28, Last annotation update)
DE Juvenile hormone esterase (EC 3.1.1.59) (JH esterase) (Fragment).
OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC Insecta; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
OC Diptera; Sphingidae; Sphingidae; Sphinginae; Manduca.
CC NCBI_TaxID=7130;
CC [1]
CC RN SEQUENCE.
CC RC TISSUE=Larval plasma;
RX MEDLINE=91072375; PubMed=2254326;
RA Venkatesh K., Abdel-Aal Y.A.I., Armstrong F.B., Roe R.M.;
RT "Characterization of affinity-purified juvenile hormone esterase from
RT the plasma of the tobacco hornworm, Manduca sexta."
RL J. Biol. Chem. 265:21727-21732(1990).
CC
CC -1- FUNCTION: JH ESTERASE PLAYS A CRUCIAL ROLE IN THE DECREASE OF
CC JH ACTIVITY IN LEPIDOPTERAN INSECTS, BY HYDROLYZING THE METHYL
CC ESTER OF JH. IT IS ALSO INVOLVED IN THE TRANSPORT OF JH.
CC -1- CATALYTIC ACTIVITY: Methyl (2E,6E)-(10R,11S)-10,11-epoxy-3,7,11-
CC trimethyltrideca-2,6-dienoate + H(2)O = (2E,6E)-(10R,11S)-10,11-
CC epoxy-3,7,11-trimethyltrideca-2,6-dienoate + methanol.
CC -1- SIMILARITY: BELONGS TO THE TYPE-B CARBOXYLESTERASE/LIPASE FAMILY.
CC PIR; A36527; A36527.
CC DR InterPro: IPR002018; Carboxylesterase.
CC DR PROSITE; PS00122; CARBOXYLESTERASE_B_1; PARTIAL.
CC DR PROSITE; PS00941; CARBOXYLESTERASE_B_2; PARTIAL.
CC KM Hydrolyase; Serine esterase.
CC NON TER 15
CC SQ SEQUENCE 15 AA; 1659 MW; D321EA4325B8848 CRC64;

Query Match 2.8%; Score 6; DB 1; Length 15;

```

Best Local Similarity 100.0%; Pred. No. 10;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 205 TEEVVV 210  
|||||  
Db 5 TEEVVV 10

## RESULT 14

UCRH\_SOLITU STANDARD; PRT; 68 AA.

AC P48504;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Ubiquinol-cytochrome c reductase complex 7.8 kDa protein (EC 1.10.2.2)  
DE (Mitochondrial hinge protein) (CR7).  
OS Solanum tuberosum (Potato).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.  
NCBI\_TaxID=4113;  
[1]  
RN SEQUENCE FROM N.A., AND SEQUENCE OF 1-18.  
RP STRAIN=cv. Desiree; TISSUE=tuber;  
RC MEDLINE=94283637; PubMed=8013669;  
RX Braun H.-P., Jaensch L., Kruft V., Schmitz U.K.;  
RT "The 'Hinge' protein of cytochrome c reductase from potato lacks the  
RT acidic domain and has no cleavable presequence.";  
RL FEBS Lett. 347:90-94(1994).  
[2]  
RN SEQUENCE OF 1-17; 43-67 AND 62-68.  
RC STRAIN=cv. Hansa; TISSUE=tuber;  
RX MEDLINE=94198758; PubMed=7764624;  
RT Braun H.-P., Kruft V., Schmitz U.K.;  
RT "Molecular identification of the ten subunits of cytochrome-c  
RT reductase from potato mitochondria.";  
RL Planta 193:99-106(1994).  
[1]  
RN FUNCTION: THIS IS A COMPONENT OF THE UBIQUINOL-CYTOCHROME C  
CC REDUCTASE COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH  
CC IS PART OF THE MITOCHONDRIAL RESPIRATORY CHAIN. THIS PROTEIN MAY  
CC MEDIATE FORMATION OF THE COMPLEX BETWEEN CYTOCHROMES C AND C1.  
CC -1- CATALYTIC ACTIVITY: O(2) + 2 ferriCytochrome c = O + 2  
CC ferriCytochrome c.  
CC -1- SUBUNIT: PLANTS BCI COMPLEX CONTAINS 10 SUBUNITS; 3 RESPIRATORY  
CC SUBUNITS, 2 CORE PROTEINS AND 5 LOW-MOLECULAR WEIGHT PROTEINS.  
CC -1- SUBCELLULAR LOCATION: Mitochondrial inner membrane.  
CC -1- DOMAIN: LACKS THE ACIDIC DOMAIN OF OTHER UCRH.  
CC -1- SIMILARITY: BELONGS TO THE UCRH/OCR6 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC  
CC EMBL: X79273; CAA55860.1; -  
CC InterPro: IPR003422; UCR\_hinge.  
DR Pfam: PF02320; UCR\_hinge; 1.  
KW Mitochondrion; Inner membrane; Electron transport; Respiratory chain;  
KW Oxidoreductase.  
FT INIT\_MET 0  
FT SEQUENCE 68 AA; 7846 MW; 2CAE7D9E5B1C6383 CRC64;

Query Match 2.8%; Score 6; DB 1; Length 68;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 CTKRVE 12  
|||||  
Db 30 CTKRVE 35

## RESULT 15

RPOL\_SULTO STANDARD; PRT; 93 AA.

ID RPOL\_SULTO  
AC Q96YA7;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE DNA-directed RNA polymerase subunit L (EC 2.7.7.6).  
GN RPOL OR STS238.  
OS Sulfolobus tokodaii.  
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;  
OC Sulfolobus.  
NCBI\_TaxID=111955;  
[1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=JCM 10545 / 7;  
RC MEDLINE=21456156; PubMed=11572479;  
RX Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,  
RA Sekine M., Baba S.-I., Ankaï A., Kosugi H., Hosoyama A., Fukui S.,  
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,  
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushiida N., Oguchi A.,  
RA Aoki K.-I., Masuda S., Yamagi M., Nishimura M., Yamagishi A.,  
RA Oshima T., Kikuchi H.;  
RT "Complete genome sequence of an aerobic thermacidophilic  
RT Crenarchaeon, Sulfolobus tokodaii strain7.";  
RL DNA Res. 8:123-140(2001).  
[1]  
RN FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription  
CC of DNA into RNA using the four ribonucleoside triphosphates as  
CC substrates.  
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
CC {RNA} (N).  
CC -1- SIMILARITY: BELONGS TO THE ARCHAEABACTERIA RPOL / EUKARYOTIC RPB1/  
CC RPB19 RNA POLYMERASE SUBUNIT FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC  
CC EMBL: AP000989; BAB67370.1; -  
CC InterPro: IPR001306; RNA\_pol\_L.  
DR Pfam: PF01193; RNA\_pol\_L; 1.  
DR ProDom: PD004240; RNA\_pol\_L\_13KD; 1.  
DR PROSITE: PS01154; RNA\_POL\_L\_13KD; 1.  
KW Transference; Transcription; DNA-directed RNA polymerase;  
KW Complete proteome.  
FT SEQUENCE 93 AA; 10454 MW; 671760C9D5F32A24 CRC64;

Query Match 2.8%; Score 6; DB 1; Length 93;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 135 EYKSI 140  
|||||  
Db 85 EYKSI 90

## RESULT 16

YE26\_AQUAE STANDARD; PRT; 100 AA.

ID YE26\_AQUAE  
AC O67419;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein AQ\_1426.  
GN AQ\_1426.  
OS Aquifex aeolicus.  
OC Bacteria; Aquificae; Aquificae (class); Aquificales; Aquificaceae;



OC Aquifex.  
 ON NCBI\_TaxID=63363;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VF5;  
 RX MEDLINE=98196666; PubMed=9537320;  
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,  
 RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aujay M., Huber R.,  
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;  
 RT "The complete genome of the hyperthermophilic bacterium Aquifex  
 aeolicus.";  
 RL Nature 392:353-358(1998).  
 CC -----  
 CC CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AB000740; AAC07390.1;  
 KM Hypothetical protein; Coiled coil; Complete proteome.  
 FT DOMAIN 14 47 COILED COIL (POTENTIAL).  
 SQ SEQUENCE 100 AA; 11858 MW; 59D1B376D128AD68 CRC64;  
 Query Match 2.8%; Score 6; DB 1; Length 100;  
 Best Local Similarity 100.0%; Pred. No. 63;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 135 EEVSKI 140  
 DB 80 EEVSKI 85  
 RESULT 17  
 GLRX\_RABIT  
 ID GLRX\_RABIT STANDARD; PRT; 106 AA.  
 AC P12864;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Glutaredoxin (Thioltransferase) (Trase).  
 GN GLRX OR GRX.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 ON NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Bone marrow;  
 RX MEDLINE=90062176; PubMed=2684977;  
 RA Hopper S., Johnson R.S., Vach J.E., Biemann K.;  
 RT "Glutaredoxin from rabbit bone marrow. Purification,  
 RT characterization, and amino acid sequence determined by tandem mass  
 RT spectrometry.";  
 RL J. Biol. Chem. 264:20438-20447(1989).  
 CC -1- FUNCTION: HAS A GLUTATHIONE-DISULFIDE OXIDOREDUCTASE ACTIVITY IN  
 CC THE PRESENCE OF NADPH AND GLUTATHIONE REDUCTASE. REDUCES LOW  
 CC MOLECULAR WEIGHT DISULFIDES AND PROTEINS.  
 CC -1- SIMILARITY: BELONGS TO THE GLUTAREDOXIN FAMILY.  
 CC PIR: A32682; GDBR.  
 DR HSSP: P35754; 1JHB.  
 DR InterPro: IPR002109; Glutaredoxin.  
 DR InterPro: IPR000063; ThioRed.  
 DR Pfam: PF00462; glutaredoxin; 1.  
 DR PRINTS: PR00160; GLUTAREDOXIN.  
 DR PROSITE: PS00195; GLUTAREDOXIN; 1.  
 KM Redox-active center: Electron transport; Acetylation.  
 FT MOD\_RES 1 1 ACETYLATION.  
 FT DISULFID 22 25 REDOX-ACTIVE.  
 FT DISULFID 78 82

SQ SEQUENCE 106 AA; 11822 MW; FEALD7CC096A2105 CRC64;  
 Query Match 2.8%; Score 6; DB 1; Length 106;  
 Best Local Similarity 100.0%; Pred. No. 66;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 190 QEFVNS 195  
 DB 2 QEFVNS 7  
 RESULT 18  
 FOLB\_STAUV  
 ID FOLB\_STAUV STANDARD; PRT; 121 AA.  
 AC P56740;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Dihydroneopterin aldolase (EC 4.1.2.25) (DHNA).  
 GN FOLB.  
 OS Staphylococcus aureus.  
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 ON NCBI\_TaxID=1280;  
 RN [1]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
 RC STRAIN=ATCC 25923;  
 RX MEDLINE=98246408; PubMed=9586996;  
 RA Hennig M., D'Arcy A., Hampele I.C., Page M.G., Oefner C., Dale G.E.;  
 RT "Crystal structure and reaction mechanism of 7,8-dihydroneopterin  
 RT aldolase from Staphylococcus aureus.";  
 RL Nat. Struct. Biol. 5:357-362(1998).  
 CC -1- FUNCTION: CATALYZES THE CONVERSION OF 7,8-DIHYDRONEOPTERIN TO 6-  
 CC HYDROXYMETHYL-7,8-DIHYDROPTERIN.  
 CC -1- CATALYTIC ACTIVITY: 2-amino-4-hydroxy-6-(D-erythro-1,2,3-  
 CC trihydroxypropyl)-7,8-dihydropteridine = 2-amino-4-hydroxy-6-  
 CC hydroxymethyl-7,8-dihydropteridine + glycolaldehyde.  
 CC -1- PATHWAY: Folate biosynthesis.  
 CC -1- SUBUNIT: HOMOOCTAMER. FOUR MOLECULES ASSEMBLE INTO A RING, AND TWO  
 CC RINGS COME TOGETHER TO GIVE A CYLINDER WITH A HOLE OF AT LEAST 13  
 CC A DIAMETER.  
 CC -1- SIMILARITY: BELONGS TO THE DHNA FAMILY.  
 CC PDB: 1DHN; 20-APR-99.  
 DR PDB; 2DHN; 20-APR-99.  
 DR InterPro: IPR003098; FOLB.  
 DR InterPro: IPR02152; FOLB; 1.  
 DR Pfam: PF00525; FOLB; 1.  
 DR TIGRFAMs: TIGR00525; FOLB\_dom; 1.  
 DR TIGRFAMs: TIGR00526; FOLB\_dom; 1.  
 KM Lyase; Folate biosynthesis; 3D-structure.  
 SQ SEQUENCE 121 AA; 13751 MW; 55B267FB9CA3D8D CRC64;  
 Query Match 2.8%; Score 6; DB 1; Length 121;  
 Best Local Similarity 100.0%; Pred. No. 75;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 135 EEVSKI 140  
 DB 59 EEVSKI 64  
 RESULT 19  
 PRO2\_DICDI  
 ID PRO2\_DICDI STANDARD; PRT; 124 AA.  
 AC P26200;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Profilin II.  
 GN PROB.  
 OS Dictyostelium discoideum (Slime mold).  
 OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.  
 ON NCBI\_TaxID=44689;  
 RN [1]  
 RP SEQUENCE FROM N.A.



RC STRAIN=AX2;  
RX MEDLINE=92226170; PubMed=1725525;  
RA Hargwitz M., Noegel A.A., Rieger D., Lottspeich F., Schleicher M.;  
RT "Dictyostelium discoideum contains two profilin isoforms that differ  
in structure and function.";  
RL J. Cell Sci. 100:481-489(1991).  
CC -1- FUNCTION: BINDS TO ACTIN AND AFFECTS THE STRUCTURE OF THE  
CYOSKELETON. AT HIGH CONCENTRATIONS, PROFILIN PREVENTS THE  
POLYMERIZATION OF ACTIN, WHEREAS IT ENHANCES IT AT LOW  
CONCENTRATIONS. BY BINDING TO PIP2, IT INHIBITS THE FORMATION OF  
IP3 AND DG.  
CC -1- SUBUNIT: OCCURS IN MANY KINDS OF CELLS AS A COMPLEX WITH MONOMERIC  
ACTIN IN A 1:1 RATIO.  
CC -1- SIMILARITY: BELONGS TO THE PROFILIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X61580; CAA43780.1; -.  
DR PIR; S18028; FADQ2.  
DR HSSP; P07763; 1ACF.  
DR DictyDB; DD01036; PROB.  
DR InterPro; IPR002097; Profilin.  
DR Pfam; PF00235; Profilin; 1.  
DR SMART; SM00392; PROF; 1.  
DR PROSITE; PS00414; PROFILIN; 1.  
KW Actin-binding; Cyoskeleton; Multigene family.  
SQ SEQUENCE 124 AA; 12729 MW; 04F5617BF0DB5A8F CRC64;  
  
OY 158 AEGKAI 163  
| | | | |  
Db 39 AEGKAI 44  
  
RESULT 20  
ID\_NUSB\_CAMJE STANDARD; PRT; 132 AA.  
AC Q9PTC0;  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE N utilization substance protein B homolog (NusB protein).  
GN NUSB OR C10382C.  
OS Campylobacter jejuni.  
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;  
OC Campylobacter.  
OX NCBI\_Taxid=197;  
| |  
RN SEQUENCE FROM N.A.  
RC STRAIN=NCTC 11168;  
RX MEDLINE=20150912; PubMed=10688204;  
RA Parthill J., Wren B.W., Mungall K., Kettley J.M., Churcher C.,  
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,  
RA Jags K., Kallingshev A.V., Moule S., Pallen M.J., Penn C.W.,  
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,  
RA Whitehead S., Barrett B.G.;  
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni  
RT reveals hypervariable sequences.";  
RL Nature 403:665-668(2000).  
CC -1- FUNCTION: Involved in the transcription termination process (By  
CC similarity).  
CC -1- SIMILARITY: BELONGS TO THE NUSB FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AL139075; CAB74218.1; -.  
DR HSSP; P04381; 1EX1.  
DR InterPro; IPR000139; NusB.  
DR Pfam; PF01029; NusB; 1.  
KW Transcription termination; Complete proteome.  
SQ SEQUENCE 132 AA; 14930 MW; 521AD68C9E07A113 CRC64;  
  
OY 36 IELAKE 41  
| | | | |  
Db 105 IELAKE 110  
  
RESULT 21  
ID\_NHNL\_BACSU STANDARD; PRT; 132 AA.  
AC P54162;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE 14.7 kDa ribonuclease H-like protein.  
GN RNHA.  
OS Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_Taxid=1423;  
| |  
RN SEQUENCE FROM N.A.  
RC STRAIN=168 / Marburg;  
RX MEDLINE=96349105; PubMed=8760912;  
RA Sorokin A.V., Azevedo V., Zumstein E., Galleron N., Ehrlich S.D.,  
RA Serror P.;  
RT "Sequence analysis of the Bacillus subtilis chromosome region between  
RT the serA and kds loci cloned in a yeast artificial chromosome.";  
RL Microbiology 142:2005-2016(1996).  
RN [2]  
RN SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=98044033; PubMed=9384377;  
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
RA Azevedo V., Bertero M.G., Bessieres P., Bolydin A., Borchert S.,  
RA Borries R., Bourcier L., Brans A., Braun M., Brignell S.C., Bron S.,  
RA Brouillet S., Brunsch C.V., Caldwell B., Capuano V., Carter N.M.,  
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
RA Denzot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,  
RA Eutian K.D., Eyringon J., Fabret C., Ferrari A., Foulger D.,  
RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizzi A., Galleron N.,  
RA Ghim S.Y., Glaser P., Goffeau A., Goughly E.J., Grandi G.,  
RA Guisepi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaat A.,  
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
RA Joris B., Karamata D., Kasahara Y., Klaerr-Bianchini M., Klein C.,  
RA Kobayashi Y., Koester P., Koningsstein G., Krogh S., Kumano M.,  
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,  
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
RA Parro V., Pohl T.M., Portecelle D., Porwollik S., Prescott A.M.,  
RA Preecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,  
RA Sato T., Scanlan E., Schleicher S., Schroeder R., Scofield F.,  
RA Sekiguchi J., Sekowska A., Serror S.J., Serror P., Shin B.S.,  
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
RA Takuchi M., Tamakoshi A., Tanaka T., Terepstra P., Tognoni K.,  
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,

RA Viari A., Wambutt R., Medler E., Medler H., Weitzenecker T.,  
 RA Winters P., Wipst A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,  
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*  
 RT *subtilis* ";  
 RL Nature 390:249-256(1997).  
 RN [3]  
 RP CHARACTERIZATION.  
 RX MEDLINE=99105810; PubMed=9888800;  
 RA Ohtani N., Haruki M., Morikawa M., Crouch R.J., Itaya M., Kanaya S.,  
 RT "Identification of the genes encoding M<sub>2</sub>+-dependent RNase HII and  
 RT M<sub>2</sub>g+-dependent RNase HIII from *Bacillus subtilis*: classification of  
 RT RNases H into three families";  
 RL Biochemistry 38:605-618(1999).  
 CC -1- FUNCTION: NOT KNOWN; DOES NOT HAVE RNASE H ACTIVITY.  
 CC -1- SIMILARITY: BELONGS TO THE RNASE H FAMILY. E88B SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L77246; AAA9618.1; -;  
 DR EMBL; Z99115; CAB14117.1; -;  
 DR HSP; P00647; IGOC.  
 DR Subtilist; BG11608; rnhA.  
 DR InterPro; IPR002156; RNaseH.  
 DR Pfam; PF00075; rnhaseh; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 132 AA; 14670 MW; 3D52E492B8541E15 CRC64;  
 Query Match 2.8%; Score 6; DB 1; Length 132;  
 Best Local Similarity 100.0%; Pred. No. 81;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 38 LAKEAI 43  
 Db 121 LAKEAI 126  
 RESULT 22  
 YGFX\_ECOLI  
 ID YGFX\_ECOLI STANDARD; PRT; 135 AA.  
 AC Q46824;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein ygfX.  
 DE Hypothetical protein ygfX.  
 GN YGFX OR B2896.  
 OS *Escherichia coli*.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC *Escherichia*.  
 OC *Escherichia*.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=9742617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Valdes J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of *Escherichia coli* K-12";  
 RL Science 277:1453-1474(1997).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC -----  
 DR EMBL; U28375; AAA83077.1; -;  
 DR EMBL; AE000373; AAC75934.1; -;  
 DR Ecogene; BG13074; ygfX.  
 KM Hypothetical protein; Transmembrane; Complete proteome.  
 FT TRANSMEM 16  
 FT TRANSMEM 36  
 SQ SEQUENCE 135 AA; 16064 MW; CB3D38F10A9D1D98 CRC64;  
 Query Match 2.8%; Score 6; DB 1; Length 135;  
 Best Local Similarity 100.0%; Pred. No. 83;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 117 AAVIL 122  
 Db 26 AAVIL 31  
 RESULT 23  
 RK16\_CYACA  
 ID RK16\_CYACA STANDARD; PRT; 140 AA.  
 AC O9RLT9;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Chloroplast 50S ribosomal protein L16.  
 GN RPL16.  
 OS *Candidium caldarium*.  
 OC Chloroplast.  
 OC Eukaryota; Rhodophyta; Bangiophyceae; Porphyridiales; Porphyridiaceae;  
 OC *Candidium*.  
 OX NCBI\_TaxID=2771;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RK-1;  
 RX MEDLINE=20496959; PubMed=11040290;  
 RA Gloeckner G., Rosenthal A., Valentin K.-U.;  
 RT "The structure and gene repertoire of an ancient red algal plastid  
 RT genome";  
 RL J. Mol. Evol. 51:382-390(2000).  
 CC -1- SIMILARITY: BELONGS TO THE L16P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF02186; AAF12914.1; -;  
 DR InterPro; IPR000114; Ribosomal\_L16.  
 DR Pfam; PF00252; Ribosomal\_L16; 1.  
 DR PRINTS; PR00060; RIBOSOMAL\_L16.  
 DR TIGRFAMs; TIGR01164; rpl16; 1.  
 DR PROSITE; PS00586; RIBOSOMAL\_L16\_1; 1.  
 DR PROSITE; PS00701; RIBOSOMAL\_L16\_2; 1.  
 KW Ribosomal protein; Chloroplast.  
 SQ SEQUENCE 140 AA; 15786 MW; CB748F6B0B90C56 CRC64;  
 Query Match 2.8%; Score 6; DB 1; Length 140;  
 Best Local Similarity 100.0%; Pred. No. 86;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 37 ELAKEA 42  
 Db 112 ELAKEA 117  
 RESULT 24  
 YE17\_YEAST  
 ID YE17\_YEAST STANDARD; PRT; 141 AA.  
 AC P40102;

DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Hypothetical 15.4 kDa protein in ISCI 3' region.  
GN YER187W.  
OS *Saccharomyces cerevisiae* (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SD88c / AB972;  
RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,  
RA Aviles E., Benro A., Brennan T., Carpenter J., Chen E., Cherry J.M.,  
RA Chang E., Duncan M., Guzman E., Hartzell G., Hinicke-Smith S.,  
RA Hyman R., Kayser A., Komp C., Lahekari D., Lew H., Lin D.,  
RA Mosedale D., Nakahara K., Namath A., Norgren R., Oelner F., Oh C.,  
RA Patel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,  
RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;  
RL Submitted (DEC-1994) to the EMBL/Genbank/DBJ databases.  
CC -1- SIMILARITY: TO YEAST KILLER TOXIN KHS AND TO YEAST YGL262W.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U18922; AAB64714.1; -.  
DR SGD; S0000989; YER187W.  
KW Hypothetical protein.  
SQ SEQUENCE 141 AA; 15395 MW; 670848E8A28312E CRC64;  
CY 164 IDQIVA 169  
DB 3 IDQIVA 8  
Query Match 2.8%; Score 6; DB 1; Length 141;  
Best Local Similarity 100.0%; Pred. No. 87;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
AC P45138;  
ID YHBC\_HAEIN STANDARD; PRT; 141 AA.  
AC 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Hypothetical protein H11282.  
GN H11282.  
OS Haemophilus influenzae.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Haemophilus.  
OX NCBI\_TaxID=727;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=rd / KW20 / ATCC 51907;  
RX MEDLINE=95350630; PubMed=7542800;  
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
RA Scott J.D., Shetty R., Liu L.-I., Glodek A., Keiley J.W.,  
RA Weiman J., Phillips C.A., Spriggs T., Heddlow E., Cotton M.D.,  
RA Usterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
RA Fine L.D., Fritchman J.L., Fuhmann J.L., Geoghegan N.S.M.,  
RA Gnehm J.C., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
RA Venter J.C.;  
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae  
RT rd.";  
RL Science 269:496-512(1995).

CC -1- SIMILARITY: BELONGS TO THE UFG090 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U32808; AAC22931.1; -.  
DR TIGR; H11282; -.  
DR InterPro; IPR003728; DUF150.  
DR Pfam; PF02576; DUF150; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 141 AA; 16108 MW; CC398AB847B886A3 CRC64;  
CY 199 OEOLIV 204  
DB 123 OEOLIV 128  
Query Match 2.8%; Score 6; DB 1; Length 141;  
Best Local Similarity 100.0%; Pred. No. 87;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
AC P02133; P06644;  
ID HBB2\_XENLA STANDARD; PRT; 146 AA.  
AC 21-JUL-1986 (Rel. 01, Created)  
DT 01-APR-1988 (Rel. 07, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hemoglobin beta-2 chain (minor) (Larval beta-II-globin) (B2G).  
GN HBB2.  
OS *Xenopus laevis* (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86067205; PubMed=2999708;  
RA Knoechel W., Meyerhof W., Stalder J., Weber R.;  
RT "Comparative nucleotide sequence analysis of two types of larval  
RT beta-globin mRNAs of *Xenopus laevis*.";  
RL Nucleic Acids Res. 13:7899-7908(1985).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87014375; PubMed=3020391;  
RA Meyerhof W., Koester M., Stalder J., Weber R., Knoechel W.;  
RT "Sequence analysis of the larval beta II-globin gene of *Xenopus  
RT laevis*.";  
RL Mol. Biol. Rep. 11:155-161(1986).  
CC -1- FUNCTION: INVOLVED IN OXYGEN TRANSPORT FROM THE LUNG TO THE  
CC VARIOUS PERIPHERAL TISSUES.  
CC -1- SUBUNIT: HETEROQUATERNARY OF TWO ALPHA CHAINS AND TWO BETA CHAINS.  
CC -1- TISSUE SPECIFICITY: RED BLOOD CELLS.  
CC -1- SIMILARITY: BELONGS TO THE GLOBIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X03142; CAA26914.1; -.  
DR EMBL; M21411; AAA49653.1; -.  
DR PIR; A02453; HBXL2.  
DR PIR; A24119; A24119.  
DR PIR; A54492; A54492.

DR HSSP, P02100, 1A9W.  
 DR InterPro: IPR000971; Globin.  
 DR Pfam: PF00042; Globin, 1.  
 DR PROSITE: PS01033; GLOBIN; 1.  
 KM Heme; Oxygen transport; Transport; Erythrocyte.  
 FT INIT\_MET 0  
 FT METAL 63 63 IRON (HEME DISTAL LIGAND).  
 FT METAL 92 92 IRON (HEME PROXIMAL LIGAND).  
 SQ SEQUENCE 146 AA; 15841 MW; BB53347CA9BFAEA CRC64;

Query Match 2.8%; Score 6; DB 1; Length 146;  
 Best Local Similarity 100.0%; Pred. No. 89;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 175 VLVDGL 180  
 |||||  
 Db 136 VLVDGL 141

RESULT 27  
 ID\_GLB\_PAREP STANDARD; PRT; 147 AA.  
 AC P80721;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Globin-3 (Myoglobin).  
 OS Paramphistomum epiclitum.  
 OC Eukaryote; Metazoa; Platyhelminthes; Trematoda; Digenea;  
 OC Echinozoemia; Paramphistomata; Paramphistomatoidea; Paramphistomidae;  
 OC Paramphistomum.  
 NC NCBI\_TaxID=54403;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=97160613; PubMed=9006947;  
 RA Rashid A.R., van Houtwert M.-L., Hague M., Siddiqi A.H., Lasters I.,  
 RA de Maeyer M., Giffon N., Marden M.C., Dewilde S., Clauwaert J.,  
 RA Vinogradov S.N., Moens L.;  
 RT "Trematode myoglobins, functional molecules with a distal tyrosine";  
 RJ J. Biol. Chem. 272:2992-2999(1997).  
 [2]  
 RN [2]  
 RP STRUCTURE BY NMR OF THE HEME POCKET RESIDUES.  
 RX MEDLINE=97160614; PubMed=9006948;  
 RA Zhang W., Rashid K.A., Hague M., Siddiqi A.H., Vinogradov S.N.,  
 RA Moens L., la Mar G.N.;  
 RT "Solution of 1H NMR structure of the heme cavity in the oxygen-avid  
 RT myoglobin from the trematode Paramphistomum epiclitum.";  
 RJ J. Biol. Chem. 272:3000-3006(1997).  
 CC -1- FUNCTION: OXYGEN BINDING PROTEIN.  
 CC -1- SUBUNIT: Monomer.  
 CC -1- MASS SPECTROMETRY: MW=16643.2; MW ERR=1.58; METHOD=Electrospray.  
 CC -1- MISCELLANEOUS: THIS GLOBIN LACKS ONE OF THE HEME-BINDING HISTIDINE  
 CC RESIDUES FOUND IN MOST OTHER GLOBINS (REPLACED BY A TYROSINE) BUT  
 CC HAS AN EXTREME OXYGEN-AVIDITY AND HIGH OXIDATION RESISTANCE  
 CC DESPITE THIS CHANGE.  
 CC -1- SIMILARITY: BELONGS TO THE GLOBIN FAMILY.  
 DR InterPro: IPR000971; Globin.  
 DR Pfam: PF00042; Globin, 1.  
 DR PROSITE: PS01033; GLOBIN; 1.  
 KM Heme; Oxygen transport; Transport.  
 FT METAL 98 98 IRON (HEME PROXIMAL LIGAND).  
 FT METAL 98 98 IRON (HEME PROXIMAL LIGAND).  
 SQ SEQUENCE 147 AA; 16639 MW; BA5062C05B8DEED3 CRC64;

Query Match 2.8%; Score 6; DB 1; Length 147;  
 Best Local Similarity 100.0%; Pred. No. 90;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 DAEGBA 162  
 |||||  
 Db 125 DAEGBA 130

RESULT 28

FLAG\_METVO STANDARD; PRT; 150 AA.  
 ID FLAG\_METVO  
 AC O06640;  
 DT 15-JUN-1998 (Rel. 36, Created)  
 DT 15-JUN-1998 (Rel. 36, Last sequence update)  
 DT 15-JUN-1998 (Rel. 36, Last annotation update)  
 DE Putative flagella-related protein G.  
 GN FLAG.  
 OS Methanococcus voltae.  
 OC Archaea; Buryarchaeota; Methanococci; Methanococcales;  
 OC Methanococcaceae; Methanococcus.  
 NC NCBI\_TaxID=2188;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DSM 1537 / PS;  
 RA Bayley D.P., Jarrell K.F.;  
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: STRONG, TO M. JANNASCHII FLAG.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb-eb.ch](mailto:license@isb-eb.ch)).  
 CC -----  
 DR EMBL: U97040; AAB57831.1; -.  
 KM Flagella; Transmembrane.  
 FT TRANSMEM 9 29 POTENTIAL.  
 SQ SEQUENCE 150 AA; 16174 MW; 40FAE641A0D5B82 CRC64;

Query Match 2.8%; Score 6; DB 1; Length 150;  
 Best Local Similarity 100.0%; Pred. No. 92;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 174 SVLVDS 179  
 |||||  
 Db 89 SVLVDS 94

RESULT 29  
 ID\_RL9\_STRPN STANDARD; PRT; 150 AA.  
 AC O97N63;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE 50S ribosomal protein L9.  
 GN RPLI OR SP2204.  
 OS Streptococcus pneumoniae.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 NC NCBI\_TaxID=1313;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TIGR4;  
 RX MEDLINE=21357209; PubMed=11463916;  
 RA Telletlin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,  
 RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,  
 RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,  
 RA Unayama L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,  
 RA Hotzapple E., Khouri H., Wolf A.M., Ulfersback T.R., Hansen C.L.,  
 RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,  
 RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,  
 RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;  
 RT "Complete genome sequence of a virulent isolate of Streptococcus  
 RT pneumoniae.";  
 RJ Science 293:498-506 (2001).  
 CC -1- FUNCTION: Binds to the 23S rRNA (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE L9P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AE007508; AAK76255.1; -  
DR TIGR; SP2204; -  
DR InterPro; IPR000244; Ribosomal\_L9.  
DR Pfam; PF01281; Ribosomal\_L9; 1.  
DR TIGRFAMs; TIGR00158; L9; 1.  
DR PROSITE; PS00651; RIBOSOMAL\_L9; 1.  
KW Ribosomal protein; rRNA-binding; Complete proteome.  
SQ SEQUENCE 150 AA; 16523 MW; 31EBDB26CF0662F1 CRC64;

Query Match 2.8%; Score 6; DB 1; Length 150;  
Best Local Similarity 100.0%; Pred. No. 92;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 158 AEGKAI 163  
Db 63 AEGKAI 68

## RESULT 30

YKH2\_CAEEL STANDARD; PRT; 152 AA.  
AC P34270;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Hypothetical protein C02C2.2 in chromosome III.  
CM C02C2.2.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Pelodertinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Br1stol N2;  
RX MEDLINE=94150718; PubMed=7906398;  
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,  
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,  
RA Fulton L., Gardner A., Green P., Hawkes T., Hillier L., Jier M.,  
RA Johnston L., Jones M., Kersey J., Kirsten J., Lalister N.,  
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,  
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,  
RA Sims M., Smailson N., Smith A., Smith M., Sonhammer B., Staden R.,  
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,  
RA Waterson R., Watson A., Weinstock L., Wilkinson-Sproat J.,  
RA Wohldman P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
RT elegans.";  
RL Nature 368:32-38 (1994).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; L23649; AAA27910.1; -  
DR PIR; S44740; S44740.  
DR WormPep; C02C2.2; CE00028.  
KW Hypothetical protein.  
SQ SEQUENCE 152 AA; 16606 MW; 2DEBF75533A121BB CRC64;

Query Match 2.8%; Score 6; DB 1; Length 152;  
Best Local Similarity 100.0%; Pred. No. 93;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Cy 42 AILKIN 47  
Db 134 AILKIN 139

## RESULT 31

LG1\_LUPLU STANDARD; PRT; 153 AA.  
ID LG1\_LUPLU  
AC P02239;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Leghemoglobin I.  
OS Lupinus luteus (Yellow lupine).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Genisteae; Lupinus.  
OX NCBI\_TaxID=3873;

RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87316940; PubMed=3628011;  
RA Konieczny A.;  
RT "Nucleotide sequence of lupin leghemoglobin I cDNA."  
RL Nucleic Acids Res. 15:6742-6742 (1987).

RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Ventus; TISSUE=Root nodules;  
RA Strozycski P.S.P.;  
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.  
RN [3]

RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Ventus;  
RA Strozycski P.W., Karlowski W.M., Legocki A.B.;  
RT "Yellow lupine gene coding for leghemoglobin I."  
RL (in) Plant Gene Register PGR98-017.  
RN [4]

RP TISSUE=Root nodules;  
RC Egorov T.A., Feigina M.Y., Karakov V.K., Shakhparonov M.I.,  
RA Mmaleva S.I., Ovchinnikov V.A.;  
RT "The complete amino acid sequence of the leghemoglobin I from yellow  
RT lupin root nodules."  
RL Bioorg. Khim. 2:125-128 (1976).  
RN [5]  
RP FUNCTION: Provides oxygen to the bacteroids. This role is  
RP essential for symbiotic nitrogen fixation.

CC -! SUBUNIT: MONOMER.  
CC -! TISSUE SPECIFICITY: Root nodules.  
CC -! SIMILARITY: BELONGS TO THE PLANT GLOBIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; Y00401; CAA68462.1; -  
DR EMBL; X77043; CAA54332.1; -  
DR EMBL; U50083; AAC04853.1; -  
DR PIR; A02561; GPVL.  
DR PIR; A26808; A26808.  
DR PIR; S41327; S41327.  
DR HSP; P02240; 2GDM.  
DR InterPro; IPR000971; Globin.  
DR InterPro; IPR001032; Leghaemoglobin.  
DR Pfam; PF00042; Globin; 1.  
DR PRINTS; PR00188; PLANTGLOBIN.  
DR PROSITE; PS00208; PLANTGLOBIN; 1.  
DR PROSITE; PS01033; GLOBIN; 1.  
KW Heme; Oxygen transport; Transport; Nitrogen fixation;

```

KM Multigene family.
FT INIT MET 0 0 IRON (HEME DISTAL LIGAND)
FT METAL 63 63 (BY SIMILARITY)
FT METAL 97 97 (IRON (HEME PROXIMAL LIGAND)
FT METAL 97 97 (BY SIMILARITY)
FT CONFLICT 79 79 Q -> E (IN REF. 4)
FT CONFLICT 120 120 E -> G (IN REF. 1)
SQ SEQUENCE 153 AA; 16622 MW; F130A15E282F1F5D CRC64;

Query March
Best Local Similarity 2.8%; Score 6; DB 1; Length 153;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 KEATLK 45
Db 111 KEATLK 116

RESULT 32
LGB2_LUPLU STANDARD; PRT; 153 AA.
ID LGB2_LUPLU
AC P02240;
DT 21-JUN-1986 (Rel. 01, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Leghemoglobin II.
OS Lupinus luteus (Yellow lupine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Genisteae; Lupinus.
OC NCBI_TaxID=3873;
RN [1]
RP SEQUENCE.
RC TISSUE=Root nodules;
RA Zgorov T.A., Kazakov V.K., Shakhparonov M.I., Feigina M.Y.,
RA Kostetsky P.V.;
RT "Primary structure of leghemoglobin II from the nodules of the yellow
RT lupin (Lupinus luteus L.).";
RL Bloorg. Khim. 4:476-480(1978).
RN [2]
RP SEQUENCE OF 61-153 FROM N.A.
RA MEDLINE=87315070; PubMed=2442601;
RA Koneczny A., Jensen E.O., Marcker K.A., Legocki A.B.;
RT "Molecular cloning of lupin leghemoglobin cDNA.";
RL Mol. Biol. Rep. 12:61-66(1987).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RA Arutjunian E.G., Kuranova I.P., Vainshtein B.K., Steigemann W.;
RT "X-ray structural investigation of leghemoglobin. VI. Structure of
RT acetate-ferri-leghemoglobin at a resolution of 2.0 A.";
RL Sov. Phys. Crystallogr. 25:43-52(1980).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (5.0 ANGSTROMS).
RC TISSUE=Root nodules;
RA MEDLINE=75118996; PubMed=1118009;
RA Vainshtein B.K., Hartunyan E.H., Kuranova I.P., Borisov V.V.,
RA Sosferov N.I., Pavlovsky A.G., Grebenko A.I., Konareva N.V.;
RT "Structure of leghaemoglobin from lupin root nodules at 5-A
RT resolution.";
RL Nature 254:163-164(1975).
CC -1- FUNCTION: Provides oxygen to the bacteroids. This role is
CC essential for symbiotic nitrogen fixation.
CC -1- SUBUNIT: MONOMER.
CC -1- TISSUE SPECIFICITY: Root nodules.
CC -1- SIMILARITY: BELONGS TO THE PLANT GLOBIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/)

```

```

CC or send an email to license@isb-sib.ch.
CC -----
DR EMBL: M17893; AAA33410.1; -.
DR PIR: A02562; GPYL2
DR PDB: 1LH1; 30-SEP-83.
DR PDB: 2LH1; 30-SEP-83.
DR PDB: 1LH2; 30-SEP-83.
DR PDB: 2LH2; 30-SEP-83.
DR PDB: 1LH3; 30-SEP-83.
DR PDB: 2LH3; 30-SEP-83.
DR PDB: 1GDY; 27-FEB-95.
DR PDB: 1LH5; 30-SEP-83.
DR PDB: 2LH5; 30-SEP-83.
DR PDB: 1LH6; 30-SEP-83.
DR PDB: 2LH6; 30-SEP-83.
DR PDB: 1LH7; 14-JUL-86.
DR PDB: 2LH7; 14-JUL-86.
DR PDB: 1GDI; 27-FEB-95.
DR PDB: 1GDK; 27-FEB-95.
DR PDB: 1GDL; 27-FEB-95.
DR PDB: 2GDM; 15-OCT-95.
DR InterPro: IPR000971; Globin.
DR InterPro: IPR001032; Leghaemoglobin.
DR Pfam: PF00042; Globin; 1.
DR PRINTS: PR00188; PLANTGLOBIN.
DR PROSITE: PS00208; PLANT GLOBIN; 1.
DR PROSITE: PS01033; GLOBIN; 1.
KM Heme, Oxygen transport; Transport; Nitrogen fixation;
KM 3D-structure; Multigene family.
FT METAL 63 63
FT METAL 97 97
FT CONFLICT 61 61
FT CONFLICT 79 79 Q -> AS (IN REF. 2; AAA33410).
FT CONFLICT 86 86 O -> E (IN REF. 1).
FT CONFLICT 150 150 T -> S (IN REF. 1).
FT CONFLICT 150 150 N -> D (IN REF. 1).
FT HELIX 5 19
FT TURN 20 21
FT HELIX 22 36
FT HELIX 38 43
FT TURN 45 49
FT HELIX 58 81
FT HELIX 88 99
FT TURN 100 101
FT HELIX 104 125
FT HELIX 128 152
SQ SEQUENCE 153 AA; 16652 MW; FE29A9DDEF33AFC8 CRC64;

Query March
Best Local Similarity 2.8%; Score 6; DB 1; Length 153;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 KEATLK 45
Db 111 KEATLK 116

RESULT 33
NUSB_FUSNN STANDARD; PRT; 153 AA.
ID NUSB_FUSNN
AC OGR11;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE N utilization substance protein B homolog (Nusb protein).
GN NUSB OR FN1616.
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteriia; Fusobacterium.
OC NCBI_TaxID=76856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;

```

MDLINE=2186394; PubMed=11893109;  
RA Kapactral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,  
RA Bhattacharya A., Bartman A., Gardner W., Grecklin G., Zhu L.,  
RA Vasiera O., Chu L., Kogan Y., Chaga O., Goldsmen E., Bernal A.,  
RA Larsen N., D'Souza M., Malinas T., Pusch G., Haselkorn R.,  
RA Forstein M., Kyriades N., Overbeek R.,  
RT "Genome sequence and analysis of the oral bacterium Fusobacterium  
RT nucleatum strain ATCC 25586."  
RL J. Bacteriol. 184:2005-2018(2002).  
CC -!- FUNCTION: Involved in the transcription termination process (by  
CC similarity).  
CC -!- SIMILARITY: BELONGS TO THE NUSB FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AE010469; AAL93731.1; -  
KM Transcription termination; Complete proteome.  
SQ SEQUENCE 153 AA; 17516 MW; 3DA8EBDBD2BE898 CRC64;  
Query Match 2.8%; Score 6; DB 1; Length 153;  
Best Local Similarity 100.0%; Pred. No. 93;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 36 IELAKE 41  
Db 126 IELAKE 131  
RESULT 34  
ID YBEY\_HAEIN STANDARD; PRT; 154 AA.  
AC P71335;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Hypothetical protein HI0004.  
GN HI0004.  
OS Haemophilus influenzae.  
CC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
CC Haemophilus.  
CC NCB1\_taxid=727;  
OK (1)  
RN SEQUENCE FROM N.A.  
RP STRAIN=Rd / KW20 / ATCC 51907;  
RC MEDLINE=95350630; PubMed=7542800;  
RA Fleischnann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
RA Keilavage A.R., Bult C.J., Tomb J.F., Dougherty B.A., Merrick J.M.,  
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
RA Weisman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
RA Usterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
RA Fine L.D., Fritchman J.L., Fuhrman J.L., Geoghegan N.S.M.,  
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
RA Venter J.C.;  
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae  
RT Rd."  
RL Science 269:496-512(1995).  
CC -!- SIMILARITY: BELONGS TO THE UPF0054 FAMILY. STRONG, TO E.COLI YBEY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; U32686; AAC21683.1; -  
DR TIGR; HI0004; -  
DR InterPro; IPR002036; UPF0054.  
DR Pfam; PF02130; UPF0054.1.  
DR ProDom; PD005688; UPF0054.1.  
DR TIGRFAMs; TIGR00043; UPF0054.1.  
DR PROSITE; PS01306; UPF0054.1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 154 AA; 17355 MW; C06F7AB7F9CA7AD CRC64;  
Query Match 2.8%; Score 6; DB 1; Length 154;  
Best Local Similarity 100.0%; Pred. No. 94;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 173 GSVLVD 178  
Db 2 GSVLVD 7  
RESULT 35  
ID RS5\_BUCAK STANDARD; PRT; 166 AA.  
AC P46183; Q44082;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE 30S ribosomal protein S5.  
GN RPSF.  
OS Buchnera aphidicola (subsp. Acyrthosiphon kondoi) (Acyrthosiphon  
OS kondoi symbiotic bacterium).  
CC Bacteria; Proteobacteria; gamma subdivision; Buchnera.  
CC NCB1\_taxid=42474;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Kurashiki;  
RX MEDLINE=96051390; PubMed=7584036;  
RA Abe R., Yamashita A., Isono K.;  
RT "Cloning and characterization of the ribosomal protein genes in the  
RT spc operon of a prokaryotic endosymbiont of the pea aphid,  
RT Acyrthosiphon kondoi."  
RL DNA Res. 1:103-114(1994).  
CC -!- FUNCTION: PROTEIN S5 IS IMPORTANT IN THE ASSEMBLY AND FUNCTION OF  
CC THE 30S RIBOSOMAL SUBUNIT (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE SSP FAMILY OF RIBOSOMAL PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; D31786; BAA06592.1; ALT\_SEQ.  
DR EMBL; D16555; BAA03984.1; -  
DR HSP; P02357; PKP.  
DR InterPro; IPR000463; Fatty acid BP.  
DR InterPro; IPR002132; Ribosomal L5.  
DR InterPro; IPR000851; Ribosomal S5.  
DR InterPro; IPR00324; Ribosomal S5\_C.  
DR Pfam; PF00333; Ribosomal S5; 1.  
DR Pfam; PF03719; Ribosomal S5\_C; 1.  
DR TIGRFAMs; TIGR01021; rpsE bact; 1.  
DR PROSITE; PS00585; RIBOSOMAL\_S5; 1.  
KW Ribosomal protein.  
SQ SEQUENCE 166 AA; 17367 MW; EDEEBAP9A720CB92 CRC64;  
Query Match 2.8%; Score 6; DB 1; Length 166;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 168 VALNSG 173  
Db 168 VALNSG 173







```

DE Fibroblast growth factor-4 precursor (FGF-4) (K-fibroblast growth
DE factor) (HBGF-4).
GN FGF4 OR FGF-4 OR KFGF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99296455; PubMed=2740210;
RA Dickson C.;
RT "The mouse homologue of hst/k-FGF: sequence, genome organization and
RT location relative to int-2."
RL Nucleic Acids Res. 17:4037-4045 (1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90201563; PubMed=2318343;
RA Hebert J.M., Basilico C., Goldfarb M., Haub O., Martin G.R.;
RT "Isolation of cdnas encoding four mouse FGF family members and
RT characterization of their expression patterns during embryogenesis.";
RL Dev. Biol. 118:454-463 (1990)
CC - FUNCTION: IS ESSENTIAL FOR SURVIVAL OF THE POSTIMPLANTATION MOUSE
CC EMBRYO AND AT LATER EMBRYONIC STAGES, IS AN ESSENTIAL COMPONENT OF
CC SIGNALING NETWORK REQUIRED FOR GROWTH AND PATTERNING OF THE
CC DEVELOPING LIMB.
CC - TISSUE SPECIFICITY: EXPRESSED IN THE BLASTOCYST INNER CELL MASS
CC AND LATER IN DISTINCT EMBRYONIC TISSUES.
CC - SIMILARITY: BELONGS TO THE HEPARIN-BINDING GROWTH FACTORS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X14849; CAA32967.1; -.
DR EMBL; M30642; AAB37619.1; -.
DR PIR; S04741; TVMSHS.
DR HSSD; P09038; 1BEG.
DR MGD; MGI:95518; Fgf4.
DR InterPro; IPR002209; HB/F_growthfact.
DR InterPro; IPR002348; ILL_HBGF.
DR Pfam; PF00167; FGF; 1.
DR PRINTS; PR00262; ILLHBGF.
DR PRODOM; PD000831; HB/F_growthfact; 1.
DR SMART; SM00442; FGF; 1.
DR PROSITE; PS00247; HBGF_FGF; 1.
DR PROTO-oncogene; Growth factor; Mitogen; signal.
KM SIGNAL
FT CHAIN 1 29
FT CONFLICT 167 167 FIBROBLAST GROWTH FACTOR-4.
FT SEQUENCE 202 AA; 21902 MW; 62D456231047CA31 CRC64;
SQ
Query Match 2.8%; Score 6; DB 1; Length 202;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OS Gallus gallus (chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=white leghorn; TISSUE=Spleen;
RX MEDLINE=96205950; PubMed=8631799;
RA Sick C., Schultz U., Staeheli P.;
RT "A family of genes coding for two serologically distinct chicken
RT interferons.";
RL J. Biol. Chem. 271:7635-7639 (1996).
CC - FUNCTION: HAS ANTIVIRAL ACTIVITIES.
CC - SUBCELLULAR LOCATION: Secreted (probable).
CC - SIMILARITY: BELONGS TO THE INTERFERON ALPHA, BETA AND DELTA
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X92479; CAA63217.1; -.
DR InterPro; IPR000471; Interferon_abd.
DR Pfam; PF00143; Interferon; 1.
DR PRODOM; PD000550; Interferon_abd; 1.
DR SMART; SM00076; IFabd; 1.
DR PROSITE; PS00252; INTERFERON_A_B_D; FALSE NEG.
KM Cyclokin; Antiviral; Glycoprotein; Signal; Multigene family.
FT SIGNAL 1 27
FT CHAIN 28 203
FT DISULFID 28 125 INTERFERON TYPE B.
FT DISULFID 57 164 BY SIMILARITY.
FT CARBOHYD 37 37 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 203 AA; 23686 MW; E9D84FEC7803FE7A CRC64;
Query Match 2.8%; Score 6; DB 1; Length 203;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 114 PQAAV 119
DB 60 PQAAV 65
RESULT 39
INFB CHICK STANDARD; PRT; 203 AA.
AC Q90873;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Interferon type B precursor.
GN IFNB OR IFN2.

```

```

OY 180 LQLON 185
DB 43 LQLON 48
RESULT 40
BCHJ RHOSH STANDARD; PRT; 206 AA.
AC Q925D7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Bacteriochlorophyll synthase 23 kDa chain (4-vinyl reductase).
GN BCHJ.
OS Rhodobacter sphaeroides (Rhodospseudomonas sphaeroides).
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Rhodobacter.
OX NCBI_TaxID=1063;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 17023 / 2.4.1 / NCIB 8253 / DSM 158;
RA Naylor G.W., Adlessee H.A., Gibson L.C.D., Hunter C.N.;
RT "The photosynthesis gene cluster of Rhodobacter sphaeroides.";
RL Photosyn. Res. 62:121-139 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 17023 / 2.4.1 / NCIB 8253 / DSM 158;
RX MEDLINE=20115911; PubMed=10648776;

```

RA Choudhary M., Kaplan S.,  
RT "DNA sequence analysis of the photosynthesis region of Rhodobacter  
RT sphaeroides 2.4.1.";  
RL Nucleic Acids Res. 28:862-867(2000).  
CC -1- PATHWAY: Light-independent bacteriochlorophyll biosynthesis.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AJ010302; CAB38730.1; -.  
DR EMBL; AF195122; AAF24280.1; -.  
DR InterPro: IPR004096; VAR.  
DR Pfam; PF02830; VAR; 1.  
KW Photosynthesis; Bacteriochlorophyll biosynthesis.  
SQ SEQUENCE 206 AA; 22128 MW; 26DB3675D1900D79 CRC64;

Query Match 2.8%; Score 6; DB 1; Length 206;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 28 HVAVRE 33  
|||  
Db 181 HVAVRE 186

Search completed: April 9, 2003, 14:43:32  
Job time : 46 secs



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: April 8, 2003, 03:51:43 ; Search time 2221 Seconds  
(without alignments)  
2856.556 Million cell updates/sec

Title: US-10-034-500-2  
Perfect score: 1119  
Sequence: 1 AEVTASCTKRVESYNVLVDY.....QEOILVTEVVVLGVNFAF 218

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+ p2n.model -DEV=xlh  
-O=/cgnr\_1/USPTO.spool/US10034500/runat\_02042003\_092744\_19703/app.query.fasta\_1.391  
-DB=GenEmbl -QFMT=faeCap -SUFFIX=rge -MINMATCH=0.1 -LOOPCU=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40 -MODE=LOCAL  
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10034500 @CGN 1.1 1687 @runat\_02042003\_092744\_19703 -NCPU=6 -ICPU=3  
-NO\_XLPEY -NO\_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-MARK TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_scs:\*  
12: gb\_ey:\*  
13: gb\_un:\*  
14: gb\_vt:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_scs:\*  
28: em\_un:\*

29: em\_vt:\*  
30: em\_hcg\_hum:\*  
31: em\_hcg\_inv:\*  
32: em\_hcg\_other:\*  
33: em\_hcg\_mus:\*  
34: em\_hcg\_din:\*  
35: em\_hcg\_rod:\*  
36: em\_hcg\_man:\*  
37: em\_hcg\_vtc:\*  
38: em\_sy:\*  
39: em\_hcgo\_hum:\*  
40: em\_hcgo\_mus:\*  
41: em\_hcgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1119	100.0	656	AX464446	AX464446 Sequence
2	126	11.3	12500	BSATPC	Z81356 B.subtilis
3	126	11.3	212610	BSUB0019	Z99122 Bacillus su
4	109	9.7	9035	AE010569	AE010569 Fusobacte
5	109	9.7	339681	AP003009	AP003009 Mesorhizo
6	103.5	9.2	3079	AF095584	AF095584 Sus scrofa
7	99	8.8	11040	AE009304	AE009304 Agrobacte
8	99	8.8	11573	AE008309	AE008309 Agrobacte
9	93.5	8.4	687	AX434018	AX434018 Sequence
10	93.5	8.4	2843	CGPRU28A	X79298 C.griseus r
11	93.5	8.4	2847	CGPRU28B	X79298 C.griseus r
12	93.5	8.4	8889	CEL243182	A1243182 Caenorhab
13	92.5	8.2	343550	AP003587	AP003587 Noctuid sp
14	92	8.2	2807	HSIARIH3	X67055 H.sapiens m
15	91.5	8.2	3181	AR204653	AR204653 Sequence
16	91.5	8.2	3043	AR204655	AR204655 Sequence
17	91.5	8.2	3195	AX322747	AX322747 Sequence
18	91.5	8.2	3204	AF127035	AF127035 Homo sapi
19	91.5	8.2	3221	AK000072	AK000072 Homo sapi
20	91.5	8.2	3265	AX092338	AX092338 Sequence
21	91.5	8.2	3265	AX376190	AX376190 Sequence
22	91.5	8.2	3265	AX403491	AX403491 Sequence
23	91	8.1	712	AP333967	AP333967 Marmota m
24	91	8.1	723	MM14137	Y14137 Marmota mon
25	91	8.1	779	AF082491	AF082491 Marmota m
26	91	8.1	1637	AF012910	AF012910 Marmota m
27	91	8.1	180461	AF250878	AF250878 Salmone11
28	91	8.1	218160	STYPHGM1	AL513383 Salmone11
29	90.5	8.1	2898	AF433571	AF433571 Cryptospor
30	90.5	8.1	13296	AE013262	AE013262 Methanosa
31	90.5	8.1	51158	ANGDHA	AE003851 Xylella f
32	90	8.0	2334	8 ANGDA	X16121 Aspergillus
33	90	8.0	7099	1 AE004083	AE004083 Xylella f
34	90	8.0	17041	1 AE004357	AE004357 Vidrio ch
35	90	8.0	19236	8 SC9346	Z48784 S.cerevisia
36	90	8.0	101054	2 AC114131	AC114131 Rattus no
37	90	8.0	324050	1 BL591983	BL591983 Listeria
38	89.5	8.0	2394	10 BC004727	BC004727 Mus muscu
39	89.5	8.0	197912	10 AL672001	AL672001 Mouse DNA
40	89	8.0	3429	3 AF224491	AF224491 Halocynth
41	89	8.0	8827	3 AF166888	AF166888 Caenorhab
42	89	8.0	8842	3 CEL243181	AJ243181 Caenorhab
43	89	8.0	8856	3 CEL243179	AJ243179 Caenorhab
44	89	8.0	8922	3 CEL243180	AJ243180 Caenorhab
45	89	8.0	110000	2 AL772337_2	Continuation (3 of

RESULT 1

ALIGNMENTS

LOCUS	AX464446	556 bp	DNA	linear	PAT 16-UU-2002
DEFINITION	Sequence 1 from Patent EP1219711.				
ACCESSION	AX464446				
VERSION	AX464446.1	GI:21899251			
KEYWORDS					
SOURCE					
ORGANISM	<i>Lawsonia intracellularis</i> .				
REFERENCE	1				
AUTHORS	Jacobs, A. A. and Vermeij, P.				
TITLE	<i>Lawsonia intracellularis</i> vaccine				
JOURNAL	Patent: EP 1219711-A 1 03-UU-2002;				
	Akzo Nobel N.V. (NL)				
FEATURES					
source	Location/Qualifiers				
	1..656				
	/organism="Lawsonia intracellularis"				
	/db_xref="taxon:29546"				
	<1...>654				
CDS					
	/note="unnamed protein product"				
	/codon_start=1				
	/transl_table=11				
	/protein_id="CAD42340.1"				
	/db_xref="GI:21899252"				
	/translation="AEVTAPSCITKRVESYNYLVDSGSMWGHVAVEPKLELAELAILI KINAMPKSYOGGILYTPAPYSVITIPSSNMSCYAECAVNTIKSDLIFGFLTPVGCG IKMEETVINGMPQAAVILLITDGHNNLDMNVEEKYSITQINPVCTHRVYSFADDAEG KATIDQIVALLNSGSVLVDGLQLQNPAPVCOEFVNSVFCQIQLTVEEVVLGRVNFAPF "				
BASE COUNT	196 a 101 c 136 g 223 t				
ORIGIN					
Alignment Scores:					
Pred. No.:	9e-107	Length:	656		
Score:	1119.00	Matches:	218		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	6	Gaps:	0		
US-10-034-500-2 (1-218) x AX464446 (1-656)					
QY	1	ATAGTUAATThrAlaSerCysThrLysAspValGluSerTyrAsnTyrLeuValAspTyr	20		
DB	1	GGGAGGTACGGCGGAGTGTACTAAACGTGTGAAGCATTAATATCTTGAGATTAT	60		
QY	21	SerGlySerMetMetMetLysHisValAlaValArgGluProLysIleGluLeuAlaLys	40		
DB	61	TCAAGCTCATGATGATGATGAACAATGTGTGTTAGAGACCTTAATAATGAATTAGCAAAA	120		
QY	41	GUUAlaIleLeuLysIleAsnAlaAlaMetProLysMetSerTyrGlnGlyLeuTyr	60		
DB	121	GAACTATATTAAAAATTATATCGCGCAATGCCAAATATCATATCAAGTGCATTAATAT	180		
QY	61	ThrPheAlaProTyrSerValIleIleProGlnGluSerTyrAsnSerCysValAlaGlu	80		
DB	181	ACTTTTGCACCTTATTTCTGTAATTAATTTCCCAAGTTCCTTGGAAATTCATGTGTGCCAA	240		
QY	81	CysAlaValAsnThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGly	100		
DB	241	TGTGGCGTTAAATACATTAATAGTCTGATTTAAGAAATTTTGGTCGTCTTACTCGTGGGA	300		
QY	101	AspGlyTyrLeuMetHisGluThrValIleAsnGlnMetProProGlnAlaAlaValIle	120		
DB	301	GACGGCATTAATAATGCATGAAACAGTCATTAATCAAAATGCCCTTCAGCAGCCCTTAAT	360		
QY	121	LeuLeuThrAspGlyHisAsnLeuGluLysMetAsnProValGlnGluValLysSerIle	140		
DB	361	CTTCTACTGATGATGATATTAATTTAAGGATGAATCCGTGAGGAGAAATTAATCTATA	420		
QY	141	TyrGlnThrAsnProAsnValCysPheHisValIleSerPheAlaAspAspAlaGlyIle	160		
DB	421	TATCAACAACAAATCTTAATGTGTCTTTTCAATGATGATTTTCAATTTGCAATGATGCTGAAGC	480		

Oy	161	LysalalalleapnglnllevalAlaleuhsnsergyservAlleuValaspGlYleu	180
Db	481	AAAGCATATATTGATTCATATTGTTGCACCTATATAGTGGAAGTCTTCTGTGATGCTTTA	540
Oy	181	GlnleuenglnaanProAlaValCysGlnGlnubheValaenservAlpheCysGlnGlnu	200
Db	541	CAGCTTACAAAAATCGTGTGTTGCCAAGAAATTGTATATAGTGTTTTGTCAAGAG	600
Oy	201	GlnlleuValAlThrGlnGlnuValValValleuAargGlyValaAnpheAlaphe	218
Db	601	CAAAATCTTGTACAGAAAGAGTTGTGTACTTGTGTGCGCTCACTTGCCTTC	654
RESULT 2			
BSATPC	BSATPC	12500 bp	DNA linear BCT 05-JUL-1999
LOCUS	DEFINITION	B. subtilis atpC gene.	
VERSION	ACCESSION	281356	
KEYWORDS	281356.1 GI:1648948		
	asparcyl-phosphate phosphatases; ATP synthase epsilon subunit; atpC		
	gene; MoaA-like protein; mura gene; narX gene; narX gene; rapB		
	gene; spoIID gene; UDP-N-acetylglucosamine		
	1-carboxyvinyltransferase; urea gene; urease gamma subunit; ywmB		
	gene; ywmC gene; ywmD gene; ywmE gene; ywmF gene; ywmG gene; ywmH		
SOURCE	Bacillus subtilis.		
ORGANISM	Bacillus subtilis		
	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.		
REFERENCE	1 (bases 10749 to 12188)		
AUTHORS	Lopez-Diaz, I., Clarke, S. and Mandelstam, J.		
TITLE	SpoIID operon of Bacillus subtilis: cloning and sequence		
JOURNAL	J. Gen. Microbiol. 132 (Pt 2), 341-354 (1986)		
MEDLINE	86226160		
PUBMED	3011962		
REFERENCE	2 (bases 9069 to 10202)		
AUTHORS	Perego, M., Hanstein, C., Welsh, K.M., Djavahishvili, T., Glaeser, P.		
	and Hoch, J.A.		
TITLE	Multiple protein-aspartate phosphatases provide a mechanism for the		
	integration of diverse signals in the control of development in B.		
	subtilis		
JOURNAL	Cell 79 (6), 1047-1055 (1994)		
MEDLINE	95094268		
PUBMED	8001132		
REFERENCE	3 (bases 6559 to 8921)		
AUTHORS	Glaeser, P., Danchin, A., Kunst, F., Zuber, P. and Nakano, M.M.		
TITLE	Identification and isolation of a gene required for nitrate		
	assimilation and anaerobic growth of Bacillus subtilis		
JOURNAL	J. Bacteriol. 177 (4), 1112-1115 (1995)		
MEDLINE	95164514		
PUBMED	7860592		
REFERENCE	4 (bases 1 to 12500)		
AUTHORS	Glaeser, P., Danchin, A., Kunst, F. and Mozer, I.		
TITLE	Bacillus subtilis atpC to urea chromosomal region		
JOURNAL	Unpublished		
REFERENCE	5 (bases 1 to 12500)		
AUTHORS	Glaeser, P.P.		
TITLE	Direct Submission		
JOURNAL	Submitted (25-OCT-1996) Philippe P Glaeser, Biochimie et Genetique		
	Moleculaire, Institut, Pasteur, 28 rue du Dr Roux, Paris CEDEX 15,		
	75728, France		
FEATURES			
source	location/Qualifiers		
	1..12500		
	/organism="Bacillus subtilis"		
	/strain="168"		
	/db_xref="taxon:1423"		
	1..136		
gene	/gene="atpC"		
	<1..136		
	/gene="atpC"		
	/codon_start=2		
	/transl_table=1		
	/product="ATP synthase epsilon subunit"		
	/protein_id="CAB03676.1"		
CDS			

/db\_xref="GI:1648849"  
/db\_xref="SWISS-PROT:P37812"  
/translation="DKERAEXAQRQERLNSQSDPTDIRRAELALQRALNRLDPAVK  
"  
complement(170. .643)  
/gene="ymmb"  
complement(170. .643)  
/gene="ymmb"  
/transl\_table=1  
/codon\_start=1  
/product="unknown"  
/protein\_id="CAB03677.1"  
/db\_xref="GI:1648850"  
/db\_xref="SPTREMBL:P70958"  
/translation="MNSFPQPEKQYGEPLPIPDQEWALAFYYDQSOQLBENELC  
QPNISLTYRESLEELNLYFOSIOELLADNPLPIEPEKMI SVYLIDCVIARHD  
AEWVYKPYKDGAYTIGFRRHRKSWHTMACDLYLRPKEDRLSLPGLVNS"  
complement(650. .658)  
1155. .1161  
1172. .1909  
/gene="ymmb"  
1172. .1909  
/gene="ymmb"  
/codon\_start=1  
/transl\_table=1  
/product="unknown"  
/protein\_id="CAB03678.1"  
/db\_xref="GI:1648851"  
/db\_xref="SPTREMBL:P70959"  
/translation="MKKKQVSHAIITISVMSFVAVPHTIHASELTPLAQMAGMERO  
DVSIDKWTLLAKQNLTEKEFYQVQRLQRYQYDVIARBDKMIKALGTYYDKN  
RTSFRQLVTLTKKKNPTSYLLYQMSLETPEDSWNTDYEQFERETLGIPOEKVYIPTC  
LNGHLDDNMNI VLQKSNQLNEFQARSVEHVBNFVYSISAFDMEEYIWTSGHKN  
LQILRSAGMGKRTVTVGPIVTTTX"  
1930. .1936  
1943. .3253  
/gene="mura"  
/gene="mura"  
1943. .3253  
/gene="mura"  
/codon\_start=1  
/transl\_table=1  
/product="UDP-N-acetylglucosamine  
1-carboxyvinyltransferase"  
/protein\_id="CAB03688.1"  
/db\_xref="GI:1648861"  
/db\_xref="SPTREMBL:P70965"  
/translation="MEKTIYRGGCKLNGTVKVEGARANVLPVTAASLASBEKSYCD  
VPTLSDYITINEVLKHLGADVHEFENNEVTVAASYALQTEAFETVRKRRSLVMGFL  
LARTGHARVALPGGCAIGSRPIDQHLGFEPMGEIIVNGGFI EAVKGRLOQAKITL  
DPSEVGATENLIMAAALAEGLTLENNAKPEIYDLANYINGMGKIRGAGTGITKIE  
GVEKLHGKHHIIPDRIEAGTFMVAAATEGNVAVKGAIVPEHLISLAKKEEMGVTK  
DEGGLAVIGPKELKPIDIKTMPPHGPPTMOSQOMALLRASGTSMTITVPENRPM  
HAEFRMRNDIKIEGSAVIINGPVQIGABVAATDTRAGALITLAGI VAEGRTRVE  
LKHLDRIYDPHQALALGADIERVNDESASEQENKEVSDLNA"  
3263. .3297  
3434. .3439  
3449. .4480  
/gene="spolID"  
3449. .4480  
/gene="spolID"  
/codon\_start=1  
/transl\_table=1  
/product="unknown"  
/protein\_id="CAB03679.1"  
/db\_xref="GI:1648852"  
/db\_xref="SWISS-PROT:P07372"  
/translation="MKQPAITLSVLCALITLLVPTLVI PQQHNKEAGASVESEKTAVS  
TKPSGAELTKASPIPVYRTANQSEVNIPLBEVYIGVASSEMPATFPEALKAQA  
LAATFIVRLVMSAVKPSGLVDDTOMFOVQKSAELKQMGTSYETFEKKITDA  
VASTQKILTYNNQPIEASPFSTNGTYENA EAVTSAIPIKSVKSPMDKSKSYVA  
TKPTTAEPQOKGVKLDGSSAVKRTIGTFRGHVATRAVINGKTLKRDIREKGLNS  
ADFEKRNQDITITVTRFGHGVMSQIGANFMAKEGKTVDIVKYTYQGTQISEADA  
FLNKRYMAKK"

conflict 3946  
/gene="spolID"  
/citation=[1]  
/replaces="g"  
4481. .4507  
4852. .4859  
4866. .5549  
/gene="ymmb"  
4866. .5549  
/gene="ymmb"  
/codon\_start=1  
/transl\_table=1  
/product="unknown"  
/protein\_id="CAB03680.1"  
/db\_xref="GI:1648853"  
/db\_xref="SPTREMBL:P70960"  
/translation="MKRPSLIIMTGLLPGLTSPAPAEKTEPTAKAPAVAVLLDAS  
GSMAKRIDVSKRNSAKETSKPASSLPEGTQVMSVFGSBGNKNSGVQSCAIRN  
VYGQSPNEQSFLNSLNTIGTGTWTPIAKALNEKSSFDOLDADAGEKVYLLITDGEET  
CGGNPIKTAKELQDNITVNVIGFDYKEGYKQINAIYAKVGGEYFPAYTQDKYEKIF  
TOOSLMASK"  
4866. .4934  
/gene="ymmb"  
5351. .5383  
5859. .5866  
5873. .6547  
/gene="ymmb"  
5873. .6547  
/gene="ymmb"  
/codon\_start=1  
/transl\_table=1  
/product="unknown"  
/protein\_id="CAB03681.1"  
/db\_xref="GI:1648854"  
/db\_xref="SPTREMBL:P70961"  
/translation="MKKLPAAGITGLTTSIASBSPAEKQADTNVAVLDPGSSMQ  
KTGERKIDIAKKS VKSFAELLPKDTNLMRVGHAQNNKLSGALSCSTETIYGLH  
PYSGSLPDNLSLEIKPYGTWTPIAKALADTRKEFAFPADGKNVYLLITDGEETCGDP  
AAIEIKLRASNVDITVINIIGFNDVKNEEMKQAAVAGGEYIISANGADEPQAMEKE  
AOKPTE"  
5873. .5941  
/gene="ymmb"  
6562. .6586  
6627. .6635  
6641. .6802  
/gene="ymmb"  
6641. .6802  
/gene="ymmb"  
/codon\_start=1  
/transl\_table=1  
/product="unknown"  
/protein\_id="CAB03689.1"  
/db\_xref="GI:1648862"  
/db\_xref="SPTREMBL:P70966"  
/translation="MKLFGMIPLIATVAFILGLVLLKLAAPFVSIITLLIAIVLFTV  
LKNQHNQRT"  
6813. .6821  
6829. .6945  
6829. .6945  
/gene="ymmb"  
6829. .6945  
/gene="ymmb"  
/codon\_start=1  
/transl\_table=1  
/product="unknown"  
/protein\_id="CAB03690.1"  
/db\_xref="GI:1648863"  
/db\_xref="SPTREMBL:P70967"  
/translation="MSTNPAFVNVAAHWCASHVIKGTAKCLGVLPKRRHHIL"  
7033. .7039

sig\_peptide  
terminator  
RBS  
gene  
CDS

sig\_peptide  
terminator  
RBS  
gene  
CDS

sig\_peptide  
terminator  
RBS  
gene  
CDS

Alignment Scores:  
Pred. No.: 0.0206  
Score: 126.00  
Length: 12500  
Matches: 57







QY	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
QY	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200
QY	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200
QY	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200
QY	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200
QY	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190										

IKTLYSTAPNAKAGIRIPFGHCKSTNSKDESGAGNELLYPIGIDLNAVEGKEIALEB  
 OPTWSTASKEIYGVEDIDIAJDOSEKTNILNIIYITDIELOCGAGVPAIAKOLKSEGN  
 IVLGIEIGNVDANOIRLLKQJADAGGYSSYANDANLUTGELIRINLARSQYWEUL  
 NDNLITVRKAGHNEILIFNKAYGSKITGEVDLSTALVGGIISGVNDVPAFAGIYIY  
 GKVDIRLKEISEERNKIDIAFEEBYRNKRKESESEYIAYLESRGSENVAPSTISVS  
 PRSAAYSTNSKGGTREDAKKDAEKIYAEKAAKQ"  
 1459. .2184  
 /gene="FN0577"  
 1459. .2184  
 /gene="FN0577"  
 /codon\_start=1  
 /transl\_table=1  
 /product="Hypothetical protein"  
 /protein\_id="AA194773.1"  
 /db\_xref="gi:19714080"  
 /transl\_table="MSN1EEFKUYNFPEEITAKDYKEIEKYLASYESGKEKEGFLP  
 VFLVDLVLKEFELDMEDENTNDIMDIYKSNLEKYDINAVEFLKNFOEMLDYFT  
 KKNYAVDRKKYKNLELSTLFNSSKRKSDVLAIVKSTKMPYELVGYFGMGVNDCELD  
 PAEOYAVDRKKYKGAVPVITVDEIEFYEKREVQTLTEAKKLVAEYHAFVYDIVEQ  
 CYGPEKTVDELKYNQWYFWMD"  
 2163. .4007  
 /gene="FN0578"  
 2163. .4007  
 /gene="FN0578"  
 /EC\_number="3.6.1.-"  
 /codon\_start=1  
 /transl\_table=1  
 /product="ATP-dependent DNA helicase recQ"  
 /protein\_id="AA194774.1"  
 /db\_xref="gi:19714081"  
 /transl\_table="MVFLVGLKAKAPALIKELYYGYNDPNEGOKETIIDILERENVL  
 GIMTGAACKSGCYQIPALIFEGULTIISPLSLMKDOYDLSRLGITASVYNSALSTSD  
 ENKILFEIKKGIKTLIYISPERLENNFELNFIVTKIIMVVDVAEACVOMGNFPR  
 SYLRADVPEKLSGETQIOTAFATAPKIXVDILKININFPYIYDFENRNIYF  
 KYHNSGIDKDLIDSKPFIIDYLRKHGKGIYICSTRKAVDIIYVLGDFDSVTRGI  
 YVDSGSEERKEDNOFLDDDIIEIMVATNAGMGINKSNIRVYHANIPALDESUYOE  
 AGRARDOGSAEALIIYNEKRDIORYLMEKEAGKTDRAVINKRLNPNMKIYAEEL  
 KTYREYLLKTFGEKMIINNYCYCENCKEKNIDFSCIENAKITISGRAKESLGIST  
 LANMLAGDARLMLNKGLDISTFGIMKEDKQEWIESFINMISEKYLLOSASFPAVL  
 KLGEXYQILINGNIKIVIRKEDEKIDFPYENDLPKELNSLKEISKOENLAPYIIFSD  
 MTLIEMAKKERNRMMDLKIIGIIGNQKFSYGEKFLIRIINYCMEERI"  
 4025. .8860  
 /gene="FN0579"  
 4025. .8860  
 /gene="FN0579"  
 /codon\_start=1  
 /transl\_table=1  
 /product="Hypothetical cytosolic protein"  
 /protein\_id="AA194775.1"  
 /db\_xref="gi:19714082"  
 /transl\_table="MFILISILAFVACCKDKDOEQNDTQGETSQGYDGYOQVLYVNA  
 NAKPISGSDIIVLFEELDEKNDQPKKLFVEVGLDGDITIMPINKLIKGNPKDONPPE  
 NAYVKSDIGKISGLSTLNDYIKVNLYLGEKKEPSSFVDSGNVLSPINNKIKFNSVNI  
 SKVLTEAVKVTNNITVOYLKLYSNEGYEMNELMDGDIIYITKEVEIIDSKEQDVOKV  
 SIDLSNTIDTGIYIYKLSAIGDSIDVDIAKESPNSEFGVDGYIYARAEKTIILSD  
 IGIYANNSNSKIDLKLNLANTLNPVNAKRELFINSKQNTLEEGTNSNGEYKSTNLD  
 KVFYLVKSGENFVLYLAGSKINVSDDIIGLSDGSLKYTTDKGYAPRGEINVE  
 SLIASKESKMDNQPFEYSFTGPDGSTIINNEIYVESNGYGTQIKTDMDLIGAMV  
 LSIRKGEIIEIOKPIEISKIANITLADIBDOEIKYKADPOKALIDFKAFLYLSGAVNV  
 DKGTIVNDVAVIEKDOOSKYREANPENBNYVOPNFAPETLLDGSQGVNLKLM  
 PDLTOSKNULYSTIYNVSDANGRSTEEKVFKIINREBSVQVQYQNDNETSVKIIIL  
 LNEKDSLVAGKQKYRIYLNKYNMWYDYIYNDERKSEKNEITVLLBEGITSGSSBS  
 ELAKTILKGDGTNFIIEDETEGSHSGGVFVKKFYGDGRKHGTEINLNTADKEKYNG  
 DLAKIKYSGAGVSKALVTIEKDGKIIKEKYMTLVKNDSEETVIEKDFPAAVYISV  
 POKYKDONDEPLRLYGSVTPMWDKSSMLLVQDTRTEVLPDGDILKLSNNKMKKQ  
 YVEVLVBEVGLKMTDYKKDPYKFFYERAKKLQVSDYNENSIIRISDYKARNLKTGQ  
 GEDAEEDLDESMALAEVAYKKEBQJLVEDAPFANVLTRFVAVSDKGVNNAVVKK  
 NFFGMRFLVAVASDESYSGASKEISVAPVPIYVTSAPVAVKVDKFTVPTLPIEIK  
 AIGNEIILTUNYKTNKAVNKOQKSEKVLFELEAPVTGKIDIDIFSKSYGKIEK  
 TNLNVNPNPQVNERKSIIVLEPNOEFTLSSEKDFINGSVKSLVLSYOKIGYKIEK  
 LIKSLIDVPYICLBOISSKGMAMLYIDMLTPIDEKNDKNEINTIIGKLNMYOARL  
 GAFAVPSQSEGISVTAIRFLTEAKKGIYVEPTNPEKSEKINLSIANTDVPKIEI  
 VLYLIAAAGDPPVNSMNIVDRYYKYNISVKKRMLIAYSGIEXGDFARKADKLPR

'T'PITTTTTV /Y/PMTTTCMMWY /Y/PMDVTWTOT /Y/PMDMCCCT /Y/P/CMCNCN

Kishida, Y., Kiyokawa, C., Kohara, M., Matsumoto, M., Matsuno, A.,  
Mochizuki, Y., Nakayama, S., Nakazaki, N., Shimo, S., Sugimoto, M.,  
Takeuchi, C., Yamada, A., and Tabata, S.  
Complete genome structure of the nitrogen-fixing symbiotic  
bacterium Mesorhizobium loti  
DNA Res. 7 (6), 331-338 (2000)  
21082930  
2 (bases 1 to 339681)  
Kaneko, T.  
Direct Submission  
Submitted (05-DEC-2000) Takakazu Kaneko, Kazusa DNA Research  
Institute, The First Laboratory for Plant Gene Research, Yana  
1532-3, Kiarazu, Chiba 292-0812, Japan  
(E-mail: kaneko@kazusa.or.jp)  
URL: http://www.kazusa.or.jp/rhizobase/  
Tel: 81-438-52-3935 (ex.2318), Fax: 81-438-52-3934)  
On May 11, 2001 this sequence version replaced gi:11994984.  
Location/Qualifiers  
1. 339681  
/organism="Mesorhizobium loti"  
/strain="WAF303099"  
/db\_xref="taxon:381"  
complement(82..420)  
/gene="nifH6350"  
complement(82..420)  
/note="unknown protein"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="BAB52662.1"  
/db\_xref="GI:14026064"  
/translation="MPLAFLDLIARIALPKVGLPERVNIKASGLTAMPHDGLRAEQG  
PFAVSMQCLRLRSLLVHRHPVLNEMRSIENVRTLLREACINMARREPPSPVCRSS  
NPALTPRCKK"  
447..593  
/gene="nifH6351"  
447..593  
/gene="nifH6351"  
/note="unknown protein"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="BAB52663.1"  
/db\_xref="GI:14026065"  
/translation="MHALVHADIDQDCGVMYAKLFGIDPFLFKLYAGGHSRGRSS  
SKRG"  
complement(733..1464)  
/gene="nifH6352"  
complement(733..1464)  
/gene="nifH6352"  
/codon\_start=1  
/transl\_table=11  
/product="transposase"  
/protein\_id="BAB52664.1"  
/db\_xref="GI:14026066"  
/translation="NWTMAWSEVARDPLVYCHRFPAEIIAHAVMLYFRPLSLRMVED  
MLAAGIIVITHTQIISMAKFGFHFARERIKRSAGCLGDMHLDCEVAINGKQOMLTA  
RVDDGGLKLETVLRSRAKARMLRLKLAQGRTPWITDKLSYDAARDLAMP  
GAYSDGKLVLENRAESHSOPTRRRERTMRPKSARQLQRFVSIHDPILNIFHPRNTL  
SSAQHRDLNAAQITWNTKACLAAG"  
complement(1080..1748)  
/gene="nifH6353"  
complement(1080..1748)  
/gene="nifH6353"  
/codon\_start=1  
/transl\_table=11  
/product="transposase"  
/protein\_id="BAB52665.1"  
/db\_xref="GI:14026067"  
/translation="MWEACATYKNNHRPIEIVAHAVMLYFRGLSLRDVEEMLESGEI  
VYSYTIIRIMKGSXSDVARRLRKASPKSDVWHLDEVLGHRVNRRLMKMGCPQO  
KLIVIRIYATAPRPSLSTRYGSTTAPSAACMLATCTCRPAASSSRIRIPFGGRRS  
GGISVSRSDGQPALMATIGISTNWWPNSASSGSGVPSIRTAWSKCWCQAEMPR  
RGSV"

```
gene
    complement (1766, .2200)
    /gene="m116354"
    /complement (1766, .2200)
    /gene="m116354"
    /note="unknown protein"
    /codon_start=1
    /transl_table=11
    /protein_id="BAB52666.1"
    /db_xref="GI:14026068"
    /translation="MWARHATIAAVERRTTPRGRVROAENHAPVSAKEIQAALPED
AMHVTTRREGSKASLASRFATVRVRPAHRDVTWRSLPPAEENFLIEMPGVEPTKTYML
STLPGKTRLADLVDAQKRWIERDYELKQSCCHIDGLTV"
gene
    complement (1939, .2931)
    /gene="m116355"
    /complement (1939, .2931)
    /gene="m116355"
    /codon_start=1
    /transl_table=11
    /product="transposase"
    /protein_id="BAB52667.1"
    /db_xref="GI:14026069"
    /translation="WLMHMGAMDLODGIETSESRPAAYVENTLASVGHADNVAPIKAY
CTGLLPERKSVPEPMARVEBGRVQAHHQSLHMFVAADNSDDAVLSRVEQVLPAL
BRQGIKRAMIVDDTGPKKKSHSVGAQYCGQKQKQNCQVAVSLVATQASLPJA
YQLYLPEBWANDPDRRAKTVGVENVFRTKEBIALAQIRALEAGSVGVVLADAGYG
NDTAFRTGLTKVGLTYVGVQSSIRLSPGTOPLPPKPMGSGRGPRAFGRLMRRS
OPMRKRCPCRPPTGTLGAPARRRLPHALPPCGFGRPIATGAPSPPLNGS"
gene
    complement (3398, .3616)
    /gene="m16356"
    /complement (3398, .3616)
    /gene="m16356"
    /note="unknown protein"
    /codon_start=1
    /transl_table=11
    /protein_id="BAB52668.1"
    /db_xref="GI:14026070"
    /translation="MLSGMPCRLLRITMMAGATIDLLAGQFLVAGFOIETRTARRH
RSAKRDRSPGRLRLMTNSARPNDRSL"
gene
    4196, .6634
    /gene="m1r6358"
    4196, .6634
    /gene="m1r6358"
    /note="unknown protein"
    /codon_start=1
    /transl_table=11
    /protein_id="BAB52669.1"
    /db_xref="GI:14026071"
    /translation="MADTSAPTHRQVVLVYIGDERKMTIGRSKSRIGDAGAGSSSR
SOSGLASGSGSGFSGVLEMRSEBGKSPSAGARGAHDDLGSGSGSPVPRRIPA
AAPRAAGPRAAGRRRSVSTGDEALAPVAGPSTGVTSPSSSGEVIETGRM
CSLIDEIOVLEKSLHSGKPMGPRDELQALNAGBELLECVALPKDEARSLTDD
VYHVRADYTHRADNSANELAHTI PKKBEORRKMEDIHNLVNSMTLHSGVNSKIM
KPYVACVERMDKQVORHELMRAVCATATLPSTBEMREEEAALRHSGVNSKIM
QIQSRINIAQVMMAPQALMLEAPIRELLIGDNGQTRLLGNTEGVEPFAVSTAEVLR
KEGRPLDIEHCHEVLEGVVERLADFPASHSVVDRLRDHQAAGLPLELTRLVDVAI
ASDEVTDLADLPKPPAPDLTDAAALAAVTDGTGKLVAAEASONSASAAALVLO
AGSSAAGREPQVAAVSDPAKAVLVRTDGTGKLVAAEASONSASAAALVLO
NPTTGENTKSTIVASTELLKPLPAQOQSLISRAHSGKPEVDHAGVAVVORQTOAE
MEPCIPALASRLSKGKMDVHRLTVQRDMLANVRLGAKMDECKPMINDCKRY
PEPSONHLEQLRSARELTPADPPCALNPEPSNLFETRLQPKALSNRTLSPVWHIT
TKRIYAQQLATDSAEFPAACHVKSNEORHNRQWOTSRKAKYDNAIHHGKLTTPAFCK
SLISDRADWRNCPA"
gene
    complement (6985, .7716)
    /gene="m116359"
    /complement (6985, .7716)
    /gene="m116359"
    /codon_start=1
    /transl_table=11
    /product="transposase"
    /protein_id="BAB52670.1"
    /db_xref="GI:14026072"
    /translation="MRTWAMSEVARDPLVRCRFPALITAAHVMILYFRFPLSLRWED
MIAANGIIVTHQTIKRSMAEKFGRHFAREIKRSAGCTAGDKMHLDECVAINKKQMLW
```

```
RAVDQDGLVLEIVQSRNKAACAKMLRKLKQAGTTPVMTIDKRSYDPAARDLMP
GVEHRSHKGLNRAENSHOPTRRERTRMKFESKARQLOREVSJHDPILANLHFERPNTL
SSAQRDLRNAAQIWNKIACLA"
gene
    8086, .17256
    /gene="m1r6361"
    8086, .17256
    /gene="m1r6361"
    /note="similar to shikimate kinase"
    /codon_start=1
    /transl_table=11
    /protein_id="BAB52671.1"
    /db_xref="GI:14026073"
    /translation="WTVGNRSGRKTEKVERTGEASSSSGSLAPAGASQSPDTV
EOMRSINRTRTPAVPMERGAADALHRRPGSPSGIHGTDIAAHHAAAPRARR
SPMSLEGETLAPVASOPTGVAAPICLSEGDIEKGRNDLTLVGMFNAADOP
GDEKLDOVINAGSALAHYKSLSPALARTVDNOLITCRVLYWDAVOESRLAEP
IHKIQKKAEBEATGLTOSTATPESVYEAACQWKEKVCGERGARARATASLP
SATAKMROEBEATTRLATGSMHKLVIYLSRIDLAAKIEPQASTDAPVREILME
NGRVRLLEAFSGDVPAPFLSTQNAVLSGGHPLADHCAVEGIVTERLSEFBLQYT
RTNLNDQTSADLPLELSQIVEGAMTVGHVHMLLQPPVAPLPPAGAIPAE
TAGSRVAKGTALQKAPALIASPADAGATTQIDTVKSPVAKGTIALPPNWPANLA
PPADMGAGIPADTGRKAAGAARRKRGKGPVAGSSAPGRASQVAADVSDTAP
ADKTVLSDMGTKVSAEDARASASAGAQAQKQAPSEMALTOBROYKPEDARVAET
VVERLQTOPAMRACPALDBPRPGLAPAOVEYDQTYRKLGMANOQDELAKALA
SOSHAEGQSSQSDAEVERTSHSPQDORMSTARVRGADEALGKSDYSTAPRGLA
GSRGATGPRRHSQDGHGEMGBOESLRERSMLDIEFSAVVENDDYRLOSQHRAT
FPAQERRRFDHVAQSFDDALNDLADLHKIFADIDIEFSAVVENDDYRLOSQHRAT
FLKRASEFTHFVPMWEDVRAQNTGTFACATSRKEPGQAGMCAQMAQVBR
LGRALNDESEKTLSTPLTFSRHRRLACENGMITIKLPHQBGALSELNGSLALL
VNGFSKMPBQENARKATVAVAAETRRRPAQSDSDPONTNVNGLSKPREECCCV
IDAIAGVCGRAGEKAGISGNSQAPNLVNGFSKMPBQENARKATVAVAAETRRRPA
EIVADISIGRLADEKAGRTGWYQALANLVNGFSKMPBQENARKATVAVAAETRRRPA
PVFDQALANLVNGFSKMPBQENARKATVAVAAETRRRPAQSDSDPONTNVNGLSKP
FSKLOAECRSATVAIGEVCHRRGKRSDFEVOHLANLVNGFSKMPBQENARKATV
NVAQEVLCARCEPSEPLPLHLASLVNNGFSKMPBQENARKATVAVAAETRRRPAQ
FEQELANLVNGFSKMPBQENARKATVAVAAETRRRPAQSDSDPONTNVNGLSKP
DCREMAVIANAVFRRGRSLGFGQDLANLVNGFSKMPBQENARKATVAVAAETRRR
TNPPDSSQYLANLVNGFSKMPBQENARKATVAVAAETRRRPAQSDSDPONTNVNGL
NGFSKMPBQENARKATVAVAAETRRRPAQSDSDPONTNVNGLSKPREECCCV
LEIARLQAGOCPPRHPAAGLSIANSIANGSLRSBAGSIAEALDKLHLAY
LHYASDRLEADAVGVTSLKALKAKLOLDYLGSLAGSAGLNRLAEIRAPGFABNNL
ETMGNLCALLPLARSPRKQMLHRRPALNLINDIOTIVHKITIAHLRASPABEARGP
```

Alignment Scores:

Pred. No.:	67	Length:	339681
Score:	109.00	Matches:	47
Percent Similarity:	41.97%	Conservative:	34
Best Local Similarity:	24.35%	Mismatches:	82
Query Match:	9.74%	Indels:	30
DB:	1	Gaps:	8

US-10-034-500-2 (1-216) x AP003009 (1-339681)

```
QY 17 LeuValaLeuTyrSerGlsSerMetMetCtyShiLeValaValaArgGluProIySelle 36
      ::::::::::::::::::::
Db 151676 ATTCTCGACGCTTCCGGCTCGATGCGGCACAG-----ATCGACGGCAAGCCAAAGCTC 151729
      ::::::::::::::::::::
QY 37 GluLeuAlaGluAlaIleLeuSylIleAenAlaIleMetProIySewSerTyrGln 56
      ::::::::::::::::::::
Db 151730 GAATGCGCAGGGAATCTTAAAGACCGCTTCAATCGATTCCCGCGCAGCATAGATC 151789
      ::::::::::::::::::::
QY 57 GlyGlyLeuTyrThrPheAlaProTyrSer----- 66
      ::::::::::::::::::::
Db 151790 GGC-----TTCAATGCGCTATGCGCATCGGAAAAAGCAGCTGCGACGACATC 151837.
      ::::::::::::::::::::
QY 67 -----ValIleIleProGlnGlySerTyrPanSerCysValaIaGluCysAlaValaIen 84
      ::::::::::::::::::::
Db 151838 CAGCTGATGCTGCGCGCTCGGCACGAGCGCATCGGATCCGCC----- 151891
      ::::::::::::::::::::
QY 85 ThrIleIySerAspIleuIlePheGlyArgLeuThrProValGlyIySpgIyIleIyS 104
      ::::::::::::::::::::
Db 151892 -----GACAGCTGTGAAGTTCTTGCGCAG--ACGCGGTGACGCGCAGCGCTCAAG 151939
```

Qy 105 MethIsgIuThrValIleAsnGlnMetProProlGlnAlaValIleLeuLeuThrAsp 124  
Db 151940 CAGGGCGCCGAGCGCTGATAATACCGAGCAAGGCGACCGCTCTCATTAACCGAC 151999  
Qy 125 GlyHIsaAnaenLeuGlyMetAsnProValGluGluValSerIleTyrlinThrAsn 144  
Db 152000 GGGCTTGAGACCTGGCGGCGCATCTCTGCTGGTGAAGAACTCAAGCATCCGGC 152059  
Qy 145 ProAnaValCySPheHisValSerPheAlaAspAspAla---GluGlyTyraAlaIle 163  
Db 152060 GTGCATTTCACGGCGGAGCTGCTCGGCTTCGATTGACGGCGGAGGCGAAGCAGATC 152119  
Qy 164 IleAspGlnIleValAlaLeuAsnSerGlySerValIleValAspGlyLeuGlnLeu 183  
Db 152120 -----GCCCTGCTGCCGCAAAACACCGCGGCAAGTACATC-----CAGGCTTCC 152164  
Qy 184 GlnAsnProAlaValCyGlnGluIuPheValaAsnSerVal 196  
Db 152165 GATGAAAGCGCTGACGAGACCGCTGTCGAGACCGTC 152203  
RESULT 6  
LOCUS AF095584 3079 bp mRNA linear MAM 13-OCT-2000  
DEFINITION Sus scrofa epithelial chloride channel protein (ABCC) mRNA,  
ACCESSION AF095584  
VERSION AF095584.1 GI:6002645  
KEYWORDS  
SOURCE  
ORGANISM  
Sus scrofa.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
1 (bases 1 to 3079)  
Gaapar,K.J., Racette,K.J., Gordon,J.R., Loewen,M.E. and  
Forsyth,G.W.  
Cloning a chloride conductance mediator from the apical membrane of  
porcine ileal enterocytes  
Physiol. Genomics (Online) 3 (2), 101-111 (2000)  
11015605  
2 (bases 1 to 3079)  
Gaapar,K.J., Gabriel,S.E., Racette,K.J. and Forsyth,G.W.  
Direct Submission  
Submitted (28-SEP-1998) Veterinary Physiological Sciences,  
University of Saskatchewan, 52 Campus Drive, Saskatoon, SK S7N 5B4,  
Canada  
FEATURES  
Source  
Location/Qualifiers  
1..3079  
/organism="Sus scrofa"  
/db\_xref="taxon:9823"  
/tissue\_type="ileal mucosa"  
1..3079  
/gene="ABCC"  
133..2886  
/gene="ABCC"  
/function="involved in exocrine chloride conductance"  
/note="similar to CACC; alternate epithelial chloride  
channel protein"  
/codon\_start=1  
/product="epithelial chloride channel protein"  
/protein\_id="AA00077.1"  
/db\_xref="GI:6002646"  
/translation="MGSFRSLFVLVHLLEGAQNSLIQLNGVGEIVAIIDPNV  
EDERLIONKMTKASPYLFEATEKRPYFNVAILLPAKAKAEVYKELTEYKNA  
DVVTEPNPNDPPTROMNGNCEKEKIFTPDPVAKGVLYOGQRFYENAH  
LRKVNRYNNEOKFYLSNKEQPVYCSALRGTVNLPQCGGSCVTRPGADVTL  
FOKECEPIPDQSEKASIMPAQSITVVEICEKKNHKAHPNDQKCNLRSTWEV  
QDSEDPFKTTPMTQPPAPFSLQIGRIYCLVLDKSGMTVGRLKRLNQAQKFL  
LQVTEQAWVGMVAIFDSAAVYKSELVQNSAERDALARSLPTAAGGTSICSLRSA  
FTVKKKKYPTGSEIVLLTQEDNTISACPEVONONGALIHITVALGSAAKELELSO  
MTGLQTYASDAANNGILDAFGALSSGNRAASRSIOLBEGQITLQNNEMKTVV  
DSYKQDTFLITTEKRLSLPIFFGVPSGSQSFVIGKKNKAIYQVETATVGM  
KYSIASSQTLTLTVSSRRSATLPPVTVS KKKMDTGKPPSPVVVYTKHQGLPIL

PAKATLISENGKTTLELDNGAGADATKNDGISRYRYTANDANGRYSVKVALGG  
VNPERRAPPLMSGAMYIRGMIEGELKNMPPREDINKDLOQKQVFSRTASGSFY  
ASDVPKPIEDLPKPKITDLCAGIQODNLINLTWAPDGDYHGRADRYIIRSTNI  
LDLRKENDSVQVNTDILPREANSEVFPKPEGIPFTGTDLFIIVQAVDKNLS  
EISNIAQVSLFLEPPARPPTPRPAPSLCPRIQVNSTIPGHIILKIMKMGLEQL  
SIA"

BASE COUNT 892 a 758 c 755 g 674 t  
ORIGIN

Alignment Scores:  
Pred. No.: 0.795 Length: 3079  
Score: 103.50 Matches: 50  
Percent Similarity: 38.03% Conservative: 39  
Best Local Similarity: 21.37% Mismatches: 78  
Query Match: 9.25% Indels: 67  
DB: 4 Gaps: 10

US-10-034-500-2 (1-218) x AF095584 (1-3079)

Qy 38 LeuAlaIyGluAlaIleLeuIySleAsnAlaIaMetProIyMetSerTyGlnIy 57  
Db 1063 CTCGACAACTGCGAAGCATACAGCTTGCGCGCTTAAAGCATGATCAACAGCAGC 1122  
Qy 58 GlyLeuTyThrPheAlaProTySerValIleIleProGlnIySerTyPAsnSerCys 77  
Db 1123 AAACCTTCCTTCTGACAG-----ACAGTGAGCAAGGGGCTGGTGGATG 1170  
Qy 78 ValaIaGluCySaIaValaAsnThrIleIySerAspLeu----- 90  
Db 1171 GTGGCATTTGACAGCGCTGCTATGTAAAGATGAATCTGTCAGATTAACAGCCGCG 1230  
Qy 91 -----GluIlePheGlyArgLeuThrPro-----ValGly 100  
Db 1231 GAAAGACAGCGGCTTGCCCGAGCTTACCACAGAGCCTGGAGAGAGATCCATTCG 1290  
Qy 101 AspGlyIleIyMethIsgIuThrValIleAsnGlnMetProProlGln-----AlaAla 118  
Db 1291 TCGGGGCTTCGCTCGCATTTACTGTGATCAAGAAATACCGACGCGGCTGTGAG 1350  
Qy 119 ValIleLeuLeuThrAspGlyHisAsn----- 128  
Db 1351 ATCGTGTACTGACGAGGAGGAGACAAACGATTAAGCGGCTGTTCCGAGGTGAG 1410  
Qy 129 -----LeuGlyMetAsnProValGluGluValIyS 138  
Db 1411 CAAAACGAGCATCATCAACAGCTGCCCTTGCGCCCTGCGCAGCAAGAGCTGAG 1470  
Qy 139 SerIleTyrgIuThrAsnProAsnValCySPheHisValaValSerPheAlaAspAspAla 158  
Db 1471 GAGCTGTGCGAGATGACAGGTGTGTGACACATATGCT-----TCAGACAGCGC 1521  
Qy 159 GluGlyIySaIaIleIleAspGlnIleValaIleAsnSerGly----- 173  
Db 1522 GAGAACACGCGCTCATTTGAGCGTTTGGGCGCTTGTGTCGGGAGACAGACTGCCCTCC 1581  
Qy 174 -----SerValIleuValaAsp-----GlyLeuGlnLeuGlnAsnProAlaValCyS 189  
Db 1582 CAGCGCTTCATCCAGCTGAGATGTCAGGCTTTAACTCCAGAACAC----- 1629  
Qy 190 GlnGluPheValaAsnSerValPheCySglnGluGln-----Ile 202  
Db 1630 ---GAATGATGAATGAGCAGAGTGGTGTGACAGCACTGTGGAAAGACACTTGT 1686  
Qy 203 LeuValIthrGluGluValaValaValaIleuArgIyValaAsnPhe 216  
Db 1687 CTCATCACTTGAACGGAAGTCTCTTCCCAATCTCTTT 1728  
RESULT 7  
LOCUS AE009304 11040 bp DNA linear BCT 20-DEC-2001  
DEFINITION Agrobacterium tumefaciens str. C58 linear chromosome, section 74 of  
167 of the complete sequence.  
ACCESSION AE009304 AE008689

VERSION	AE009304.1	GI:17742153
KEYWORDS	Agrobacterium tumefaciens str. C58 (U. Washington).	
SOURCE	Agrobacterium tumefaciens str. C58 (U. Washington).	
ORGANISM	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group; Rhizobiaceae; Rhizobium.	
REFERENCE	1 (bases 1 to 11040)	
AUTHORS	Wood, D.W., Setubal, J.C., Kaul, R., Monks, D., Chen, L., Wood, G.E., Chen, Y., Woo, L., Kitajima, J.P., Okura, V.K., Almeida Jr., N.F., Zhou, Y., Bovee Sr., D., Chapman, P., Clendenning, J., Deatherage, G., Gillet, W., Grant, C., Guenther, D., Kutyavin, T., Levy, R., Li, M., Mclelland, E., Palmeri, A., Raymond, C., Rouse, G., Saenphimachak, C., Wu, Z., Gordon, D., Eisen, J.A., Paulsen, I., Karp, P., Romero, P., Zhang, S., Yoo, H., Tao, Y., Biddle, P., Jung, M., Krespan, W., Perry, M., Gordon-Kamm, B., Liao, L., Kim, S., Hendrick, C., Zhao, Z., Dolan, M., Tingey, S.V., Tomb, J., Gordon, M.P., Olson, M.V. and Nester, E.W.	
TITLE	The genome of the natural genetic engineer Agrobacterium tumefaciens C58	
JOURNAL	Science 294 (5550), 2317-2323 (2001)	
MEDLINE	21608550	
REFERENCE	11743193	
AUTHORS	2 (bases 1 to 11040)	
TITLE	Wood, D.W., Setubal, J.C., Kaul, R., Monks, D., Chen, L., Wood, G.E., Chen, Y., Woo, L., Kitajima, J.P., Okura, V.K., Almeida Jr., N.F., Zhou, Y., Bovee Sr., D., Chapman, P., Clendenning, J., Deatherage, G., Gillet, W., Grant, C., Guenther, D., Kutyavin, T., Levy, R., Li, M., Mclelland, E., Palmeri, A., Raymond, C., Rouse, G., Saenphimachak, C., Wu, Z., Gordon, D., Eisen, J.A., Paulsen, I., Karp, P., Romero, P., Zhang, S., Yoo, H., Tao, Y., Biddle, P., Jung, M., Krespan, W., Perry, M., Gordon-Kamm, B., Liao, L., Kim, S., Hendrick, C., Zhao, Z., Dolan, M., Tingey, S.V., Tomb, J., Gordon, M.P., Olson, M.V. and Nester, E.W.	
TITLE	Direct Submission	
JOURNAL	Submitted (27-SEP-2001) Department of Microbiology, University of Washington, 1959 NE Pacific Ave, Box 357242, Seattle, WA 98195-7242, USA	
FEATURES	Location/Qualifiers	
SOURCE	1..11040	
gene	/organism="Agrobacterium tumefaciens str. C58 (U. Washington)"	
gene	/strain="C58"	
gene	/db_xref="taxon:180835"	
gene	complement(183..1007)	
gene	/gene="Atu3730"	
gene	complement(183..1007)	
gene	/gene="Atu3730"	
gene	/note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer"	
gene	/codon_start=1	
gene	/transl_table=11	
gene	/product="conserved hypothetical protein"	
gene	/protein_id="AA144540.1"	
gene	/db_xref="GI:17742154"	
gene	/translation="MRLFLGDMVGKGTAVMDRLPGHISDLKLPVINGENAGGFCITIDIFLETINAGDVTGTGNHWDEKAEFCGRHQPFPANYPFGTGGSGI FFAARGARIVANIGRVFHPBELDPPESAETIIAAGCKQADALIFDEFAETSE KQCFRHPVDGRASFPVGTHTHTPTADQILNGSTAVMSDAGCGIDSLGMEKEPI NPFISKMPKRGREASGPATTCGVEISDRGLAEKAPRIKRGISLTPSFWA"	
gene	complement(1107..1694)	
gene	/gene="Atu3731"	
gene	complement(1107..1694)	
gene	/gene="Atu3731"	
gene	/note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer"	
gene	/codon_start=1	
gene	/transl_table=11	
gene	/product="5-formyltetrahydrofolate cyclo-ligase"	
gene	/protein_id="AA144541.1"	
gene	/db_xref="GI:17742155"	
gene	/translation="MGGDAAEKARLNGERLAAADALTPAREOQKSLMTAHGVSIGPFAGTGVSFMPTRSEVDTPLEALPARAGRLVLPVLDQETIVFVFPADPLVYTG FDTGSGDEADAVLDPHILVPLSVFDGEGRIGYGAGHYDRAIARLHAKGRQPVILGI	
gene	AFDQCEVASVPABPHVALHVLTEGRLWFSQR"	
gene	complement(1694..2956)	
gene	/gene="tlyC"	
gene	/note="Atu3732"	
gene	complement(1694..2956)	
gene	/gene="tlyC"	
gene	/note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer"	
gene	/codon_start=1	
gene	/transl_table=11	
gene	/product="hemolysin"	
gene	/protein_id="AA144542.1"	
gene	/db_xref="GI:17742156"	
gene	/translation="MIVLLTVLVNGVLAMSELAVSSRPRLKYVLAQSGRGATWALG LSENGRFLSTVQIGITLVGLVSGAFSGATGARTGMLLEQGVDPRAADAIIGVSIV VAIYLSILVIGELVPOIALRDPKRIARVAFPMVMSRIGAPLVLLDSSGKVALAI LGHSGNSAIVTDEITRTVLAHSAVITEESAMTSTVRLADRARAGIMPRRV EVDVIDESAERIRQLOKOTQSRSLPVRNGSDELICVLPKADAFDALASKEIDVRL LREPVVSDITSADVIOQLRSTVRVAVLYDEIGHFEGIVSSGVLEALITGAFQEN DEEPAMVEREDSGFLVAGMMPADEFAFVRMGFOIDEDAEFFTVAGLVDEFRRRLPELG HIRNGRFRFEVIDDGHRIKVLVSRRA"	
gene	complement(2974..3186)	
gene	/gene="Atu3733"	
gene	complement(2974..3186)	
gene	/gene="Atu3733"	
gene	/note="putative; ORF located using Glimmer"	
gene	/codon_start=1	
gene	/transl_table=11	
gene	/product="hypothetical protein"	
gene	/protein_id="AA144543.1"	
gene	/db_xref="GI:17742157"	
gene	/translation="MKFRVNPGPSPPDYFFVDSYPSLLPGMAFLYCHCNGGTIAYPD MYVCHLLHKRCGQNRDVPNSPKS"	
gene	3291..3914	
gene	/gene="Atu3734"	
gene	3291..3914	
gene	/gene="Atu3734"	
gene	/note="putative; ORF located using Glimmer"	
gene	/codon_start=1	
gene	/transl_table=11	
gene	/product="hypothetical protein"	
gene	/protein_id="AA144544.1"	
gene	/db_xref="GI:17742158"	
gene	/translation="MDRWPOGLWFLTSGASERIPDGRIGNHQRADSEFCWRGKFAF SMGRGFSCEQLFAGNAVQVQIADGNSHPIRRLVFGHAVTVWGKCNLTLPIDL AVRFKLVHHDAGDGEPOIARFSEALEMDPTTVEGRIVGQVTFPRACHAIGL AVDINSDRHRTAPYRSTARIYMAVTSRSDTSRLSSRLANNSA"	
gene	3923..4195	
gene	/gene="Atu3735"	
gene	3923..4195	
gene	/gene="Atu3735"	
gene	/note="putative; ORF located using Glimmer"	
gene	/codon_start=1	
gene	/transl_table=11	
gene	/product="hypothetical protein"	
gene	/protein_id="AA144545.1"	
gene	/db_xref="GI:17742159"	
gene	/translation="MCEBGAVMDPDLIDILARPLPLDTHIDIGEPICDSTQLRQ RRLGLSRHRFQIARIPRNPVSGLPNNHEVITRTIPSIKPVH"	
gene	4408..6390	
gene	/gene="tKta"	
gene	/note="Atu3736"	
gene	4408..6390	
gene	/gene="tKta"	
gene	/note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer"	
gene	/codon_start=1	
gene	/transl_table=11	
gene	/product="transketolase"	
gene	/protein_id="AA144546.1"	
gene	/db_xref="GI:17742160"	
gene	/translation="MISDRKIRNANAIFLAMDAVEKANGHGLPMGADVATVLF TRYLFKDPKAPLMADRDVFLSAGHSMFLVSLVLTGYEDMTIDEIKRFRFGSKTA	

## gene

## CDS

gene  
CDS

GHPEYGHATGIETTGTGGLGCIANAVGMAIERKLEEFSGDLSQSHFTYVLGCGGLM  
EGISHAIALAGHKLKNTLVFMDNNNTTIGEVGLSTQIARFQAVHNTTRVDS  
HDPDAIAAIAEAAQSDRPTIACKTVGFQAPKQGHKNGHNPGLAEI;AAARSL  
NMEBAFVPIEDVDAMRLAGLRSTKQDMDWARLEATETAKKEFRFGDLPGNF  
DSIDAFKKTIENNPTVATRKASDESLEVINGLIPENVGSSADLTBSNATKQMS  
ITTPDFSGRTYHYGIREHGMAMNGILHGGFLFYAGGFLFYSCYCPST;RLAALMG  
IRVAVLTHDSIGVEEDGPTHQPEQIALALAIPLVLRADDETETMECQALBOK  
HLRPSALATRONLTPSRKYESEKNSAYGAYELVASADAKYSIFASGSEVVALKAA  
TLKDQSVARVSVCFELPKQEPYTNALIGKAPVKIAVEAARIQGMDFYIGDGA  
FVGMHSGASAPADOLFPHGFGITAEVVAVAEAKLA"

/gene="gapa"  
/note="Atu3737"  
6453..7463  
/gene="gapa"  
/note="gapa"  
/note="identified by sequence similarity; putative; ORF  
located using Blastx/Glimmer"  
/codon\_start=1  
/transl\_table=1  
/product="Glyceroldehyde 3-Phosphate Dehydrogenase"  
/protein\_id="AAI44547.1"  
/db\_xref="GI:17742161"

/translation="MTVVAINGRGRIGRNVRAIVESGRDIEVVAINDGPETNA  
HLRYDSIHGFPAVDKVEGDTTIYVGGKPIKTVARPDATLPHKELGVDIATLCTGI  
FTARDKAAHLTAGAKRVIVAPADGDLTVFVNDQTLKHLVLSNACCTNCLV  
PVVVLDDVIGIDHGFMTTISHTYDQDPTLDMHDLRAAALSMIPSTGAAKAV  
GLVPLHKGKLDGTSIRVPTPNVSVVDKFTAKRPTVEEVNVAIKSANGKLGIIIG  
YTEPELVSRDNDHSHSIFAIDQTKMEGNFVRLVSNYDNEMGNSNMSDTVAFAK  
TI"

7780..9645  
/gene="Atu3738"  
7780..9645  
/gene="Atu3738"  
/note="identified by sequence similarity; putative; ORF  
located using Blastx/Glimmer"  
/codon\_start=1  
/transl\_table=1  
/product="Na+/H+ antiporter"  
/protein\_id="AAL44548.1"  
/db\_xref="GI:17742162"

/translation="MESFYLLLVATVVLVAARSSLIARFGAPLLILFVIGLAAG  
VDGIEFTNPPLVAMGSLILAVILFDSGTSIQSRILSAAPVVALATGVLITSV  
FPAGAAALLGLSMLEGLLGAIVATAAAPFLIRGGHIRQVNSTLEVSNGA  
DPMALFLVALVEIVAKQGLAGIDSGFLIFIEBMGLGVFGLLGGLMILTVNRFA  
ADRGAPFLVALALVFSFGAIGSGFLVYVAGIVAGNRIPAKXTIRFHEGLT  
MLAOLIMLMGLIATSPGPPAIPALVILFIVARPLAVLMSLPRYTOQETS  
FVAVVGLRGAVSILALNPIIGGLDDAOITRNAFTIVLVLVGMTIKRVATRLGL  
IVPRMGVDKLEVDLPETANHELISYVAGSALMOGERIPRAMPSELVVDGKSIR  
YVAGRLENDLVILFIAPNYTRLIDRLFAFALPVADDADPFGITISPERPAKEME  
AAYPGILISPAEHMTIETRIELRAGAGVADRVLQPLVILVALDEGRATIGVGI

## Alignment Scores:

Pred. No.: 11.1 Length: 11040  
Score: 99.00 Matches: 46  
Percent Similarity: 37.87% Conservative: 18  
Best Local Similarity: 27.22% Mismatches: 63  
Query Match: 8.85% Indels: 42  
DB: 1 Gaps: 7

US-10-034-500-2 (1-218) x AEO09304 (1-11040)

Qy 65 TyrSerValIleIleProGlnGlySerTPanSerCyValAlaGluCyAlaValaIen 84  
Db 9668 CATAGTATGTCCTCCGCACTCCGCAACAAATCAGATGACCCATGCCCTTTCAAG 9727  
Qy 85 ThrIleYSerAspIeuGlnIlePheGlyArgLeu----- 96  
Db 9728 ACCATCGACGACCTTAACGACATGCGGAAAGCGCTTCTCGTCGCGCTGATCTCAAC 9787  
Qy 97 ThrProValIleYAspGlyIleYLeuMetHisGluThrValIleAsnGlnMetProProGln 116  
Db 9788 GTGCGGTCGCGGATGGCAAGGTGACGACGACCCGATGCAAGCGCTTGGCGGACC 9847

Qy 117 -----AlaIleValIleLeuLeuThrAspGlyHisAsnAsn 128  
Db 9848 ATTCTGAACCTCTCTTCAAGAGCGCAAAAGTCATTCGTCGCCATTCGGCGGCCCA 9907  
Qy 129 LeuGlyMetAsnProValGlnGlu-----ValYSerIleTyGlnThrAsnProAsn 146  
Db 9908 AAGGCT---GAACCCGTCGTGAGATGTCGCTTCCAGANTGTCGCCAGCTCGAGAC 9964  
Qy 147 ValCyPheHisValIleValSerPheAlaIleAspAlaGlnGlyLeuAlaIleLeuAspGln 166  
Db 9965 GTCTGATCATCCGCCATTCCTTCGCAACGACATCGATCGGCCCCCTCGACGGATGCG 10024  
Qy 167 ILeValAlaLeuAsnSerGlySerValIleValAspGlyLeuGlnLeuGlnAsn--- 185  
Db 10025 GTCCGAAAATGATGATGATGCGCATATCTCG-----CTCGGAAAACACCC 10069  
Qy 186 -----ProAlaValCyGlnGln----- 191  
Db 10070 CGCTTCCACAGGCGGAGAGAAAGACATCCCGCTTCGTGAGAACTGGCGCGCAT 10129  
Qy 192 -----PheValAsnSerValPhe 197  
Db 10130 GGCGATATCTATGTGAACGACGCTTC 10156

RESULT 8  
AE008309/c 11573 bp DNA linear BCT 18-DEC-2001  
LOCUS Agrobacterium tumefaciens str. C58 linear chromosome, section 113  
DEFINITION AE008309 AE007870  
ACCESSION AE008309.1 GI:15159562  
VERSION AE008309.1 GI:15159562  
KEYWORDS  
SOURCE  
ORGANISM Agrobacterium tumefaciens str. C58 (Cereon).  
Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
Rhizobiaceae; Rhizobium.  
REFERENCE  
AUTHORS Hinkle,G., Slater,S.C. and Goodner,B.  
TITLE Complete Genome Sequence of Agrobacterium tumefaciens C58 (Rhizobium radiobacter C58), the Causative Agent of Crown Gall Disease in Plants  
JOURNAL Unpublished  
REFERENCE  
AUTHORS Hinkle,G., Slater,S.C. and Goodner,B.  
TITLE Direct Submission  
JOURNAL Submitted (14-AUG-2001) Bioinformatics, Cereon Genomics, 45 Sidney Street, Cambridge, MA 02139, USA  
COMMENT Approximately 800 bp of telomeric sequence missing from the left end of the chromosome and 200 bp missing from the right end.  
FEATURES  
source  
1..11573  
/organism="Agrobacterium tumefaciens str. C58 (Cereon)"  
/strain="C58"  
/db\_xref="taxon:181661"  
75..1718  
/gene="AGR\_L\_2178"  
75..1718  
/gene="AGR\_L\_2178"  
/note="conserved hypothetical protein - Deinococcus radiodurans (strain R1)"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="AGR\_L\_2178p"  
/db\_xref="GI:15159563"  
/translation="MERYDVIYVGAGLAGLVAATEAERGSVCVVDQEGRONLGGQA  
FWSLGLFVDSPEQRRMVARSDSLARDFSGAGPDRPDHPRRAEYLDPLFAG  
EKRMELRMGRMRPVPVGMAREGSLADGHSVPRPHVMTGTGVLAPVKKAEEM  
AASRLTFPRRHVDRLDTDGRTGTSGLVLAIRGCRGSRRLKCDPRPSASAI  
IVSGGIGIGNOELTRRMPPVPRDLRGPENNVCPYAPVVDGRMGTITLACGTAVNBR  
MWHYTBGVKNHDPVWNHGRIRILPFGSSFWCDAGNLDAPMGFGLTKLNGER  
GSSHWFILTKAIIKKEPPLSGSQRNDLTGKVDRLILKRGKPPQPVRAFMERGED  
FVVDLTLEELVAEMNAIGDRNLNIDHTRQILEADREIENGFSKDAQVTALHGARVYL



gene	CDS	gene	CDS
<p> <code>/gene="AGR_L.2190"</code>  <code>/note=" (AC006919) putative fructose biphosphate aldolase"</code>  <code>/codon_start=1</code>  <code>/transl_table=11</code>  <code>/product="AGR_L.2190p"</code>  <code>/protein_id="AK89667.1"</code>  <code>/db_xref="GI:15159569"</code> </p>	<p> <code>/complement (8219...9499)</code>  <code>/gene="AGR_L.2193"</code>  <code>/note=" (AF256214) phosphoglycerate kinase; Pgk"</code>  <code>/codon_start=1</code>  <code>/transl_table=11</code>  <code>/product="AGR_L.2193p"</code>  <code>/protein_id="AK89667.1"</code>  <code>/db_xref="GI:15159569"</code>  <code>/translation="MRPDRLSAISRHSWVRPELATNODRMPAFKTIIDDLNAGKVV LAVVDLNVPAVDKVDYDRIERVAPTITELSSKAKVLLIARFGKGPVAVEMSL QIIVPYEDVLDHAISFATDTCIGPADAVAKMDGDIILLENTRFHGEKNDPAPE ELANGDIYVNDAFSAHRAHASTEGIARLPVYAGTQWAELEALKGQVPRPVQ AIVAGKVSCKIDLLMNLVKKVDALVYGGMATFPLAARTGVNKGSCEDLAEATQ IMEAATSGCAIVLPEGVVAREFTGAANEIVDINALPADAVLIDVGPKSVESIKAM ISAEETLVMNGLPAGPEIPEPDAATVAAAKHAECTTAGCLVSAGGDTVVALNHAG VSDDFYITISAGAEFLEMGEKLPGVAILITPAK"</code>  <code>/complement (9489...11417)</code>  <code>/gene="AGR_L.2194"</code>  <code>/complement (9489...11417)</code>  <code>/gene="AGR_L.2194"</code>  <code>/note="probable sodium/hydrogen antiporter Pao21 (imported) - Pseudomonas aeruginosa (strain PAO1)"</code>  <code>/codon_start=1</code>  <code>/transl_table=11</code>  <code>/product="AGR_L.2194p"</code>  <code>/protein_id="AK89668.1"</code>  <code>/db_xref="GI:15159570"</code>  <code>/translation="MTADVDSIRCGGRGKYEAVDVSFYLILVATVLYVAASFSL FARFAGPALLLLPLVIGLAGVGLGIEFTNNPLAWLSLILAVLIFDSGFSTLOS IRISAPGAPVALATGVGILVSVFAGAAALGLISWEGELIGAVTSDAAVAFELIR IGGHIIDVRSITLVEVSGNDPMALPLTVALVEIVAGGGLAGISGFFLIFEENG LGVFLIGLGLMIAIVNRRPADRGALPIVLAALLVFSGTAGIGSGFLAVVAGI VAGNRRIFAKETIRRHGEGTWTLAQIIMPTMGLIATPSPGPALAPVLAFLFIVP ARPLAVLSTMPENYTOQETSFAVWGLRAVSLILAIMPLIGLDDAQYFNNAPIII VLVLAVQWTKIKVATRIQLIVPRRGVDDKLEVDLPGTANHELISRYAKSGALFIQ GERIPRWAMPSLVVRDCKSIRYOYAGLRENDIVLYFIANRYRLIDRLPASALPVAD DDDFEPIFTISPRPAKEMAAVGPGLIPSEHAMITAILIERLKGAGYADVRLV GPLVILRALDEQALIVGGLISLEPVRPALGLPIFISFDILRRAPFHIAQRDLRSAA DAGGVSVAAPKTYTENNA"</code>  <code>/db_xref="GI:15159570"</code> </p>	<p> <code>/complement (8219...9499)</code>  <code>/gene="AGR_L.2193"</code>  <code>/complement (8219...9499)</code>  <code>/gene="AGR_L.2193"</code>  <code>/note=" (AF256214) phosphoglycerate kinase; Pgk"</code>  <code>/codon_start=1</code>  <code>/transl_table=11</code>  <code>/product="AGR_L.2193p"</code>  <code>/protein_id="AK89667.1"</code>  <code>/db_xref="GI:15159569"</code>  <code>/translation="MRPDRLSAISRHSWVRPELATNODRMPAFKTIIDDLNAGKVV LAVVDLNVPAVDKVDYDRIERVAPTITELSSKAKVLLIARFGKGPVAVEMSL QIIVPYEDVLDHAISFATDTCIGPADAVAKMDGDIILLENTRFHGEKNDPAPE ELANGDIYVNDAFSAHRAHASTEGIARLPVYAGTQWAELEALKGQVPRPVQ AIVAGKVSCKIDLLMNLVKKVDALVYGGMATFPLAARTGVNKGSCEDLAEATQ IMEAATSGCAIVLPEGVVAREFTGAANEIVDINALPADAVLIDVGPKSVESIKAM ISAEETLVMNGLPAGPEIPEPDAATVAAAKHAECTTAGCLVSAGGDTVVALNHAG VSDDFYITISAGAEFLEMGEKLPGVAILITPAK"</code>  <code>/complement (9489...11417)</code>  <code>/gene="AGR_L.2194"</code>  <code>/complement (9489...11417)</code>  <code>/gene="AGR_L.2194"</code>  <code>/note="probable sodium/hydrogen antiporter Pao21 (imported) - Pseudomonas aeruginosa (strain PAO1)"</code>  <code>/codon_start=1</code>  <code>/transl_table=11</code>  <code>/product="AGR_L.2194p"</code>  <code>/protein_id="AK89668.1"</code>  <code>/db_xref="GI:15159570"</code>  <code>/translation="MTADVDSIRCGGRGKYEAVDVSFYLILVATVLYVAASFSL FARFAGPALLLLPLVIGLAGVGLGIEFTNNPLAWLSLILAVLIFDSGFSTLOS IRISAPGAPVALATGVGILVSVFAGAAALGLISWEGELIGAVTSDAAVAFELIR IGGHIIDVRSITLVEVSGNDPMALPLTVALVEIVAGGGLAGISGFFLIFEENG LGVFLIGLGLMIAIVNRRPADRGALPIVLAALLVFSGTAGIGSGFLAVVAGI VAGNRRIFAKETIRRHGEGTWTLAQIIMPTMGLIATPSPGPALAPVLAFLFIVP ARPLAVLSTMPENYTOQETSFAVWGLRAVSLILAIMPLIGLDDAQYFNNAPIII VLVLAVQWTKIKVATRIQLIVPRRGVDDKLEVDLPGTANHELISRYAKSGALFIQ GERIPRWAMPSLVVRDCKSIRYOYAGLRENDIVLYFIANRYRLIDRLPASALPVAD DDDFEPIFTISPRPAKEMAAVGPGLIPSEHAMITAILIERLKGAGYADVRLV GPLVILRALDEQALIVGGLISLEPVRPALGLPIFISFDILRRAPFHIAQRDLRSAA DAGGVSVAAPKTYTENNA"</code>  <code>/db_xref="GI:15159570"</code> </p>	<p> <code>/complement (8219...9499)</code>  <code>/gene="AGR_L.2193"</code>  <code>/complement (8219...9499)</code>  <code>/gene="AGR_L.2193"</code>  <code>/note=" (AF256214) phosphoglycerate kinase; Pgk"</code>  <code>/codon_start=1</code>  <code>/transl_table=11</code>  <code>/product="AGR_L.2193p"</code>  <code>/protein_id="AK89667.1"</code>  <code>/db_xref="GI:15159569"</code>  <code>/translation="MRPDRLSAISRHSWVRPELATNODRMPAFKTIIDDLNAGKVV LAVVDLNVPAVDKVDYDRIERVAPTITELSSKAKVLLIARFGKGPVAVEMSL QIIVPYEDVLDHAISFATDTCIGPADAVAKMDGDIILLENTRFHGEKNDPAPE ELANGDIYVNDAFSAHRAHASTEGIARLPVYAGTQWAELEALKGQVPRPVQ AIVAGKVSCKIDLLMNLVKKVDALVYGGMATFPLAARTGVNKGSCEDLAEATQ IMEAATSGCAIVLPEGVVAREFTGAANEIVDINALPADAVLIDVGPKSVESIKAM ISAEETLVMNGLPAGPEIPEPDAATVAAAKHAECTTAGCLVSAGGDTVVALNHAG VSDDFYITISAGAEFLEMGEKLPGVAILITPAK"</code>  <code>/complement (9489...11417)</code>  <code>/gene="AGR_L.2194"</code>  <code>/complement (9489...11417)</code>  <code>/gene="AGR_L.2194"</code>  <code>/note="probable sodium/hydrogen antiporter Pao21 (imported) - Pseudomonas aeruginosa (strain PAO1)"</code>  <code>/codon_start=1</code>  <code>/transl_table=11</code>  <code>/product="AGR_L.2194p"</code>  <code>/protein_id="AK89668.1"</code>  <code>/db_xref="GI:15159570"</code>  <code>/translation="MTADVDSIRCGGRGKYEAVDVSFYLILVATVLYVAASFSL FARFAGPALLLLPLVIGLAGVGLGIEFTNNPLAWLSLILAVLIFDSGFSTLOS IRISAPGAPVALATGVGILVSVFAGAAALGLISWEGELIGAVTSDAAVAFELIR IGGHIIDVRSITLVEVSGNDPMALPLTVALVEIVAGGGLAGISGFFLIFEENG LGVFLIGLGLMIAIVNRRPADRGALPIVLAALLVFSGTAGIGSGFLAVVAGI VAGNRRIFAKETIRRHGEGTWTLAQIIMPTMGLIATPSPGPALAPVLAFLFIVP ARPLAVLSTMPENYTOQETSFAVWGLRAVSLILAIMPLIGLDDAQYFNNAPIII VLVLAVQWTKIKVATRIQLIVPRRGVDDKLEVDLPGTANHELISRYAKSGALFIQ GERIPRWAMPSLVVRDCKSIRYOYAGLRENDIVLYFIANRYRLIDRLPASALPVAD DDDFEPIFTISPRPAKEMAAVGPGLIPSEHAMITAILIERLKGAGYADVRLV GPLVILRALDEQALIVGGLISLEPVRPALGLPIFISFDILRRAPFHIAQRDLRSAA DAGGVSVAAPKTYTENNA"</code>  <code>/db_xref="GI:15159570"</code> </p>

```
Db 9406 ACCATCGACGACCTTAACACATTCGCCGAAAGCGCTTTCGCGCGCTTGAATCTCAAC 5347
Qy 97 ThrProValGlyAspGlyIleuYmechIsgIuThValIleasnImeCProProGln 116
Db 9346 GCGCCGCTGCGGATGCGAAGGTAACGACGACCGCATGCAAGCGCTTCGCCGAC 5287
Qy 117 -----AlaAlaValIleLeuLeuThraSpGlyIhIasnaen 128
Db 9286 ATTCGGAACCTCTTCGAAAGGCGCAAAAGTCATTCCTGCCCATTCGCGCCCA 5227
Qy 129 LeuGlymeAsnProValGluGlu-----ValIyserIleTyrgIlnThraSnProaen 146
Db 9226 AAGGCT---GAACCCGCTGCTAGATGTCGCTTCCAGATTGTGCCACGCTGCAAGAC 5170
Qy 147 ValCyPhehIleValIseSerPheAlaAspAlaGluGlyValAlaIleIleAspGln 166
Db 9169 GTCGCGATCAGACGCAATTCCTTCGCAACGACGATGCGTCCCTGACGCGATGCG 5110
Qy 167 ILeValAlaLeuAsnSerGlySerValIleuValAspGlyLeuGlnIleuLeuGlnasn--- 185
Db 9109 GTGCGAAAATGATGATGCGATATTCG-----CTGCTGGAACAAACACC 5065
Qy 186 -----ProAlaValCyseGlnGlu----- 191
Db 9064 CGCTTCCACAGGCGGAGAAAGAACGATCCCGCTTCGTGAAAGAACTGCGCGCAT 5005
Qy 192 -----PheValAsnSerValPhe 197
Db 9004 GCGCATATCTATGTGAACGACGCTTC 8978

RESULT 9
AX434018 687 bp DNA linear PAT 28-JUN-2002
LOCUS AX434018
DEFINITION Sequence 2433 from Patent WO0229113.
ACCESSION AX434018
VERSION AX434018.1 GI:21658826
KEYWORDS
SOURCE
ORGANISM Bacillus licheniformis.
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE
1 Berka, R. and Clausen, I. G.
TITLE Methods for monitoring multiple gene expression
JOURNAL Patent: WO 0229113-A 2433 11-APR-2002;
Novozymes Biotech, Inc. (US) ; Novozymes A/S (DK)
FEATURES
source
1. 687
/organism="Bacillus licheniformis"
/db_xref="taxon:1402"
BASE COUNT 218 a 147 c 177 g 145 t
ORIGIN

Alignment Scores:
Pred. No.: 1.38 Length: 687
Score: 93.50 Matches: 46
Percent Similarity: 34.43% Conservative: 27
Best Local Similarity: 21.70% Mismatches: 84
Query Match: 8.36% Indels: 55
DB: Gaps: 8

US-10-034-500-2 (1-218) x AX434018 (1-687)
Qy 17 LeuValAspTyrsSerGlySerMetMetIysHisValAlaValArgIuProIysIle 36
Db 124 ATGCTGATGCAAGCGGCGCATGCTAAAGATCGCGGGGCTC-----TCAAATATAC 177
Qy 37 GluLeuAlaIysGluAlaIleLeuYsIle-----AsnAlaAla 49
Db 178 GAGCTTGCCAAATGAAGCGTTCAAGCTTGCAAGCTTGAATGCAAAACGTCCT 237
Qy 50 MetProIysMetSerTyrgIlnGly----- 57
```

```
Db 228 ATGAGATTTTCGCTTCAGAGGAAACATATAAATCCGAAAGTCACTGTCAT 297
Qy 58 -----GlyLeuTyThraPheAlaProIysSerValIleIleProGlnGlySerTrp 74
Db 298 GCATCAGAGGGCTTACGGCTTCGAAACGATAC----- 333
Qy 75 AsnSerCyValAlaGluCyAlaValaIleThraIleYsSerAspLeuGluIlePheGly 94
Db 334 -----GAGCAAAAGCTTCGCCCAATTCCTTAAACGCGCATCGCA 369
Qy 95 ArgLeu-----ThraProValGlyAspGlyIleuYmechIsgIuThValIleasnIln 112
Db 370 CCGACCGGATGACACCGCATGCAAAAGCGCTGCAAGATGCAAAACGCGTTACACG 429
Qy 113 MetProProGlnAla-----AlaValIleLeuLeuThraSpGlyIhIasnaenIleuGly 130
Db 430 CTGCAACAACAGGAAAGAAAGCTGCTATCTGCTGACAGACGGTGAAGAAACATCGGA 489
Qy 131 MetAsnProValGluGluValIysSerIleTyrgIlnThraSnProaenValCysePhehIis 150
Db 490 GGCATTCGCGTAAAGTGCACAAACAGAACTGCGCAAAATCCAAATGCGGTTGTC-----AAC 543
Qy 151 ValValAserPhehIleAspAlaGluGlyValAlaIleIleAspGlnIleValAlaLeu 170
Db 544 GTGATCGGCTTTGATTATGAAGAGACTTCCACGCAATTCACGATGCGACGACT 603
Qy 171 AsnSerGlySerValIleuValAspGlyLeuGlnIleuLeuGlnAsnProAlaValCyseGln 190
Db 604 GCGCGCGGTGAATATTC-----CAGCAAAAACATAA 636
Qy 191 GluPheValAsnSerValPheCyseGlnIleuIle 202
Db 637 AATGACATCAAAAGAAATTTTACTCAGAGACCAATT 672

RESULT 10
CGPRU28B 2843 bp DNA linear ROD 15-SEP-1994
LOCUS CGPRU28B
DEFINITION C. griseus repetitive DNA, clone pRU 2.8b.
ACCESSION X79299
VERSION X79299.1 GI:587100
KEYWORDS
SOURCE
ORGANISM Chinese hamster.
Cricetulus griseus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
Cricetulus.
REFERENCE
1 (bases 1 to 2843)
AUTHORS Fatyol, K., Cserpan, I., Praznovsky, T., Kereso, J. and Hadlaczky, G.
TITLE Cloning and molecular characterization of a novel chromosome
JOURNAL specific centromere sequence of Chinese hamster
MEDLINE Nucleic Acids Res. 22 (18), 3728-3736 (1994)
PUBMED 7937084
REFERENCE
2 (bases 1 to 2843)
AUTHORS Fatyol, K.
TITLE Direct Submision
JOURNAL Submitted (16-MAY-1994) K. Fatyol, Institute of Genetics,
Biological Research Center of the, Hungarian Academy of Sciences,
6701 Szeged, POB 521, HUNGARY
FEATURES
source
1. 2843
/organism="Cricetulus griseus"
/db_xref="taxon:10029"
/chromosome="2"
/clone="pRU2.8b"
/clone_1lb="CHO-K1 genomic library"
/note="centromere"
1. .2843
/note="tandem repeat"
/evidence=experimental
BASE COUNT 718 a 751 c 731 g 643 t
ORIGIN
```



Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
7.83	93.50	42.05%	23.86%	8.36%	2843	42	32	59	43	8

US-10-034-500-2 (1-218) x CGPRU28B (1-2843)

```

Qy 5 AlasercyThlyAryValGluSerTyraenTyrlauValAspTySerGlySerMet 24
Db 76 GCCGCTGCACGGGTAGAGTTCCGAGTCCAGATTCCTT-----TCCTCG 120
Qy 25 MetMetLySHiVaValAlaValArgIuProLySileGluLeuAlaValaIleLeu 44
Db 121 CCTGAGCCCAATGAGGAGTGCACGCGACGAGACTGAGCGTGCACAGACAGCTTTG 180
Qy 45 LySileAsnAlaAlaMetPro-----51
Db 181 CCACGTGCTGCTGCATCCTGTGGGACAGGACCTTGACCATGATGCCCATCAACC 240
Qy 52 -----LysMetSerTyrlnglyLeuTyrlThPheAlaProTySerValIlelle 69
Db 241 TCCTACCGAATGCTCGGCTGCA-----GAGACGCGACAGCCCTCTCCACCAAGCTC 294
Qy 70 ProGlnGly-----SerTPaSerCySValAlaGluCysAlaValaAnthrile 86
Db 295 CAGAGACACACACCGCTTGATGAAACACCTGCGTGGGACGCTCTTTCTTATGACAGT 354
Qy 87 LySeraSpleuGluIlePheGlyArgLeuThrProValGlyAspGlyIleLySMeHis 106
Db 355 ATGAAGAAGATCCAAAGATGACCAAGATCTCCGACCCCTGACGAGACTCTGATCCCTAC 414
Qy 107 GluThrValIleAsnGlnMetProPro-----115
Db 415 GATCAA--TGGGAGACAGATGACGACCAACCTTGAGCCCTCTGACGAGAAAGTCTAT 471
Qy 116 GlnAlaAlaValIleLeuLeuThraSpGlyHis-----AsnAsnLeuGly 130
Db 472 CAAGCTCTACGTGATGACCTCTGTGACACACTGCAAGGCTTACGTCATCAACGGA 531
Qy 131 MetAsnProValGluGluVallySeriLeTyrlGlnThraSProAsn 146
Db 532 TGGCCGCTCTTCCAGAAAGAACACCTT--ACACCAAAACCCCAAC 576

RESULT 11
LOCUS CGPRU28A 2847 bp DNA linear ROD 15-SEP-1994
DEFINITION C.griseus repetitive DNA, clone prU 2.8a.
ACCESSION X79298
VERSION X79298.1 GI:587099
KEYWORDS centromeric region; repetitive DNA; satellite-like sequence.
SOURCE Chinese hamster.
ORGANISM Cricetus griseus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
Cricetulus.
1 (bases 1 to 2847)
REFERENCE
AUTHORS Fatyol,K., Caserpan,I., Praznovazky,T., Kereso,J. and Hadlaczky,G.
TITLE Cloning and molecular characterization of a novel chromosome
specific centromere sequence of Chinese hamster
JOURNAL Nucleic Acids Res. 22 (18), 3728-3736 (1994)
MEDLINE 95023119
PUBMED 7937084
REFERENCE
AUTHORS Fatyol,K.
TITLE Direct Submission
JOURNAL Submitted (16-MAY-1994) K. Fatyol, Institute of Genetics,
Biological Research Center of the, Hungarian Academy of Sciences,
6701 Szeged, POB 521, HUNGARY
FEATURES
source 1..2847

```

```

/organism="Cricetus griseus"
/db_xref="taxon:10029"
/chromosome="2"
/clone="prU2.8a"
/clone_lib="CHO-K1 genomic library"
/note="centromere"
repeat_region
1..2847
/note="tandem repeat"
/evidence=experimental
BASE COUNT 711 a 755 c 729 g 652 t
ORIGIN

```

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
7.84	93.50	42.05%	23.86%	8.36%	2847	42	32	59	43	8

US-10-034-500-2 (1-218) x CGPRU28A (1-2847)

```

Qy 5 AlasercyThlyAryValGluSerTyraenTyrlauValAspTySerGlySerMet 24
Db 76 GCCGCTGCACGGGTAGAGTTCCGAGTCCAGATTCCTT-----TCCTCG 120
Qy 25 MetMetLySHiVaValAlaValArgIuProLySileGluLeuAlaValaIleLeu 44
Db 121 CCTGAGCCCAATGAGGAGTGCACGCGACGAGACTGAGCGTGCACAGACAGCTTTG 180
Qy 45 LySileAsnAlaAlaMetPro-----51
Db 181 CCACGTGCTGCTGCATCCTGTGGGACAGGACCTTGACCATGATGCCCATCAACC 240
Qy 52 -----LysMetSerTyrlnglyLeuTyrlThPheAlaProTySerValIlelle 69
Db 241 TCCTACCGAATGCTCGGCTGCA-----GAGACGCGACAGCCCTCTCCACCAAGCTC 294
Qy 70 ProGlnGly-----SerTPaSerCySValAlaGluCysAlaValaAnthrile 86
Db 295 CAGAGACACACACCGCTTGATGAAACACCTGCGTGGGACGCTCTTTCTTATGACAGT 354
Qy 87 LySeraSpleuGluIlePheGlyArgLeuThrProValGlyAspGlyIleLySMeHis 106
Db 355 ATGAAGAAGATCCAAAGATGACCAAGATCTCCGACCCCTGACGAGACTCTGATCCCTAC 414
Qy 107 GluThrValIleAsnGlnMetProPro-----115
Db 415 GATCAA--TGGGAGACAGATGACGACCAACCTTGAGCCCTCTGACGAGAAAGTCTAT 471
Qy 116 GlnAlaAlaValIleLeuLeuThraSpGlyHis-----AsnAsnLeuGly 130
Db 472 CAAGCTCTACGTGATGACCTCTGTGACACACTGCAAGGCTTACGTCATCAACGGA 531
Qy 131 MetAsnProValGluGluVallySeriLeTyrlGlnThraSProAsn 146
Db 532 TGGCCGCTCTTCCAGAAAGAACACCTT--ACACCAAAACCCCAAC 576

RESULT 12
LOCUS CEL243182 8889 bp mRNA linear INV 16-DEC-1999
DEFINITION Caenorhabditis elegans mRNA for inositol 1,4,5-trisphosphate
receptor (itr-1 gene), splice variant 4.
ACCESSION AJ243182
VERSION AJ243182.1 GI:5262810
KEYWORDS inositol 1,4,5-trisphosphate receptor; itr-1 gene; splice variant.
SOURCE Caenorhabditis elegans.
ORGANISM Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditioidea; Rhabditidae; Peloderinae; Caenorhabditis.
1 (bases 1 to 8889)
REFERENCE
AUTHORS Baylis,H.A., Furutachi,T., Yoshikawa,F., Mikoshiba,K. and
Satellie,D.B.

```

TITLE Inositol 1,4,5-trisphosphate receptors are strongly expressed in the nervous system, pharynx, intestine, gonad and excretory cell of *Caenorhabditis elegans* and are encoded by a single gene (*itr-1*)

JOURNAL J. Mol. Biol. 294 (2), 467-476 (1999)

MEDLINE 20079510

REFERENCE PUBMED 10610772

AUTHORS 2 (bases 1 to 8889)

TITLE Baylis, H.A.

JOURNAL Direct Submission

Submitted (15-JUN-1999) Baylis H.A., Department of zoology, University of Cambridge, Downing Street, Cambridge, CB1 3EU, UNITED KINGDOM

FEATURES

source Location/Qualifiers

gene 1. .8889 /organism="Caenorhabditis elegans"

5'UTR 1. .88 /isolate="Bristol N2"

misc\_RNA 1. .8889 /db\_xref="taxon:6239"

3'UTR 1. .8889 /def\_stage="mixed"

1. .8889 /country="United Kingdom"

1. .8889 /gene="itr-1"

1. .8889 /gene="itr-1"

1. .8889 /note="trans-spliced leader SL1"

89. .8662 /gene="itr-1"

89. .8662 /function="intracellular IP3 gated calcium channel"

89. .8662 /note="splice variant 4: exons (1., 2., 3+) (10., 11+) (23+) (29-)"

89. .8662 /codon\_start=1

89. .8662 /product="inositol 1,4,5-trisphosphate receptor"

89. .8662 /protein\_id="CAB45863.1"

89. .8662 /db\_xref="GI:5262811"

89. .8662 /db\_xref="SPTREMBL:Q9Y0A0"

89. .8662 /translation="MDIDRRKRVLEIRNIGALPSRKLTIANSDHGNNGNIHIGDIIS LTTSSSSNQEQRGFLSTGLVDRCIVELKDGRESPEKRRDCIFKCPNRAAOK HLMTQKRFQGTGDMFDDLNKLKVAADKESESEFQCTLGNVIOYGMOLLYV KSNKYITVOKNSPAKREBNAMKVLDKAGNENPILIEPAKVHAIQGNVAGNKISL I PNVSTTQAGHVSQSLHSFNLLDHQSAEVNCLNENPTEWQVPMFLPBNQNSV KSGDVRLPHADQOTFLITDITPKONPPTDVVFLMTRNPSADATSRALMEVQVQ TNAYRGATKANAKVREFKHLATDMYLSAPVOVQKAPAMNRPPLLSLQSGSRPME RTNRPAMAYSDGPNVNESDTTQONIPVYVVLGPTSEPEEDANLIFQDSTEM KSNKEVPRSVYRLHOSDQKVAHATNATEKONHNSKNGKGVKICEKRNVDKET PALPNVDEVRDLDPANDACKALNPKIKIGOVISKESTTOLLINICIFLVN SSDLADPILKISDPSPRDROKILREOVNOVFLIKAPLPKQGTTELPPLSSRS ELDSRNEIFKTMQOLCYCLIKTSQVSTRKQERFLAEFGELIQOIGEDMLAEITRA VLNPNPKLEKTVTPHYERFVELVRNNRQKFLDYADLCVCRGEANKKIQELICTS VLSKRDIDPMDTKIIDGELVGNAPNPKLVDAEGKXNSDDAEHLDDYRHLIDL SOMQEOQYLAIDPPERLMMNISQQLPAELVLCQMSDNRLPYDRGSPFTMLHLHY VRGSPMAIRHARLWMSIPENVNISTVESVEVADSGSRMRIGEGAKHLATVETRY LMGRNOSMEROSVNSKLTVEITVNLAKALQAFNYSFNDLLQNTONLALINBGRATEQVPSHRAMVNLRNNSKSMRGKNSKDKLAKTSVTAEEKRTREGALVAKK LIVAELIOFMDVRDRYRTVMSWFKNVPCEJDGSLMHSASINERASLTYAIRT SSGEHLHDGDLQDLALILQMTMSDVPPLCTALKFHFPOYQELIDELQVOLL VSNVDENVYRQIDRDLFILKLTKESELVHGRHNSIDKEDEKERTTEHLOVH DLKSPRAFDGSDMEALMAVLEHNPYSRNECQULNRLLIKDNRDAVAVLOLSXK APLIAYPLIRQMLVRLTGMCYRKGDPPQDTNNOOLKNNRMYEVLVPIVPHDKKD HDMMKLTITSHETPLRSFCKTKENOSLTKYKISTEKQALMEVETIEGTVVAIL RNNEELASNPFEELIAHVGILHNSRNPFLLELQALVCYDEISGGQKANEIC AASDEQVQLVYNDASFELEAMKDEKESKSSDSRKLVTYHLEVLRLAMCRKGN GNTLKCASQIIPMDHIVAVVTAKQCLVEKCVYVSLDLCYHIDDAEKQVKTREYD HILNLLIEDIRSLVEKLGTAEATLEHYICHTVEVLKFEFAPYASLQOAKYDVH HKTTFSELELELTYLBRGKLKSGSSNNWVAVACIRLTITMAEENHTTIPATLAGO MSGOTSVROKQOAASSAKVIGIKRLARQNTLNGRLVYTSNSNTBHSANVTCV HMTIGERFTHPLHABAGVYVLEVLHPELPEGSALRQCARGVVALILQCHT LMQKQNDLCARVLOTLCKMCDCTKQOLTHQGHVNSVSEKQOLGVALILQCHT LNNHNPILDRQOSKIGEVLEAVKEKEETWSEBDLYAIOCLNAGASLDTVDTIIE PSREIPLKAILHARALHLEGNDKVOHSPYMSKOKOIHPEPKAILIRIOAORLKS DMASCDSEKSVSTVLTPLIDAGTGNALFVPOOVHPSISMSQSLNDLTHET PDLAAYDSEKSTALPPEVALVETPLKAVLOLLENHNSLQNTLRKQSDTTHNALS ETLSFLDTVCGSTKSGLVGVEIGEHNSFLITQTLATLTTEFCQCPCHENDMTAMQEN

ALIGNMENT SCORES: 31.5 Pred. No.: 8889 Score: 93.50 Matches: 52 Percent Similarity: 37.13% Conservative: 36 Best Local Similarity: 21.94% Mismatches: 72 Query Match: 8.36% Indels: 77 Gaps: 12

US-10-034-500-2 (1-218) x CEL243182 (1-8889)

Qy 15 AenTyrLeuValAspTyrSerGlySerMetMetMetCysHisValAlaValArgGluPro 34

Db 5849 AATGATAGGTTCAACACTTTTTCACGCGAGGAGCAAAAGACATCATGACCA 5908

Qy 35 -----LysIleGluLeuAla-----LysGluAlaIle 43

Db 5909 TTCTTCAAAAGCTATTTTACATGAAATTCACAGCTTCAGAAACAGATTGAAAGTATATG 5968

Qy 44 LeuYsIleAsnAlaAlaMetProLysMetSer----- 54

Db 5969 ATGAGTTCAGTACAGCAAAACAAAGATTTCATCTACTGTTCTAACCTCTATGAT 6028

Qy 55 -----TyrGlnGlyGlyLeuTyrThrPheAlaProTyrSerValIleIle 69

Db 6029 GCTGGTATACGGGTTCATGAGCATATTATTGAA-----GTT 6067

Qy 70 ProGlnGlySerTyrPheAsnSerCysValAlaGluCysAlaValAsnThrIleLysSerAsp 89

Db 6068 CCACAAACAGTTCGTCATCTTCATTCGGA-----ATGTCCAATTAAACACAGAT 6121

Qy 90 Leu-----GluIlePheGlyArgLeuThrProValGlyAspGlyIleLysMetHisGluThr 108

Db 6122 TTGACACACTGATTCGATTTTGAACCTTATGCAAGT-----GAGGAAAA 6169

Qy 109 ValIleAsnGlnMetProProGlnAlaAlaValIle----- 120

Db 6170 TCCACTGATGATGCCCTCCGAAAGTGCACCTTGTGAACCAATCTTCGTTCTTCA 6229

Qy 121 LeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsnProValIleGluValIleLysSerIle 140

Db 6230 CTATTATATGAAACACACACAGCTTCTTCAAACTTCTTCGT-----AAACAATCT 6283

Qy 141 TyrGlnThrAsnProAsnValCysPheHisValValSerPheAlaAspAspAlaGluGly 160

Db 6284 GATGAAACAAATTCACAACTGTTTGAACCTTATGCTT----- 6325

Qy 161 LysAlaIleIleAspGlnIleValAlaLeuAsnSerGlySerValIleValAspGly--- 179

Db 6326 -----TTGATATACGCTGTGATCTTCAAAAGGAAAGCTTGGAGTGTGTTGGTGA 6376

OY 180 -----leuGlnleuLen 183  
 DB 6377 ATGGAGAGCACAATTCTCTATTACACAAACACTTGCCAGCTGATGATTTCTGT 6436  
 OY 184 GlnaenPcoAlaVal1CysGlnGlnPheValAaenSerValPheCysGlnGln 200  
 DB 6437 CAAGTCCCA-----TGTCAATGAGATCATGAAATACATGCGCATGCAAGAA 6481  
 RESULT 13  
 AP003587/c 343550 bp DNA linear BCT 28-NOV-2001  
 LOCUS Nostoc sp. PCC 7120 DNA, complete genome, section 7/19.  
 DEFINITION AP003587 BA000019  
 VERSION AP003587.1 GI:17130808  
 KEYWORDS  
 SOURCE Nostoc sp. PCC 7120 DNA.  
 ORGANISM Nostoc sp. PCC 7120  
 Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
 REFERENCE  
 AUTHORS 1  
 Kaneko, T., Nakamura, Y., Wolk, C.P., Kurita, T., Sasamoto, S.,  
 Watanabe, A., Itoh, S., Ishikawa, A., Kawashima, K., Kimura, T.,  
 Kishida, Y., Kohara, M., Matsumoto, M., Matsuno, A., Muraki, A.,  
 Nakazaki, N., Shimo, S., Sugimoto, M., Takazawa, M., Yamada, M.,  
 Yasuda, M. and Tabata, S.  
 Complete genomic sequence of the filamentous nitrogen-fixing  
 cyanobacterium *Anabaena* sp. strain PCC 7120  
 JOURNAL DNA Res. 8 (5), 205-213 (2001)  
 MEDLINE 21595285  
 REFERENCE 2 (bases 1 to 343550)  
 AUTHORS Kaneo, T.  
 JOURNAL Direct Submission  
 TITLE Submitted (02-MAY-2001) Takakazu Kaneko, Kazusa DNA Research  
 Institute, The First Laboratory for Plant Gene Research; Yana  
 1532-3, Kisarazu, Chiba 292-0812, Japan  
 (E-mail: kaneko@kazusa.or.jp,  
 URL: http://www.kazusa.or.jp/cyanobase/,  
 Tel: 81-438-52-3935 (ex.2338), Fax: 81-438-52-3934)  
 FEATURES  
 source location/Qualifiers  
 1..343550  
 /organism="Nostoc sp. PCC 7120"  
 /db\_xref="taxon:103690"  
 /note="synonym: *Anabaena* sp. PCC7120"  
 complement(94..411)  
 /gene="a111719"  
 complement(94..411)  
 /gene="a111719"  
 /note="ORF ID: a111719  
 hypothetical protein"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="BAB73418.1"  
 /db\_xref="GI:17130809"  
 /translation="MRKKSAILEAVHETAEADLHKAGLNNQTLREFEHLCLPPEPL  
 EFLQIKETRESSQVSAVPAIRILINISTIVQKWEIGQKPSGASLKLHLVANGRLNS  
 VLY"  
 complement(504..782)  
 /gene="a111720"  
 complement(504..782)  
 /gene="a111720"  
 /note="ORF ID: a111720  
 hypothetical protein"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="BAB73419.1"  
 /db\_xref="GI:17130810"  
 /translation="MKIKNRPEDRMARKEGLKNLSICNAVNMAGLVDADLGGSLF  
 KRIAPGKSGKSGFRTLIATNNEDVGLFLAFQKTNATVLIHMKKGL"  
 935..1507  
 /gene="a111721"  
 935..1507  
 /gene="a111721"  
 /note="ORF ID: a111721"  
 gene  
 CDS  
 gene  
 CDS

/codon\_start=1  
 /transl\_table=11  
 /product="transcriptional regulator"  
 /protein\_id="BAB73420.1"  
 /db\_xref="GI:17130811"  
 /translation="MRHKDKKQALICDAALIELITANGFADTSKSIKAAANSPATI  
 VYVPEENKELINKLYLCKOESIAEILKGVNOMISVKAEMINWNNYKAIKNPVPF  
 AFTQFANSPMVDRIKCDQESLSYKPMIDLFNRGKEKVFEDIPLEIFTAFTAPLIG  
 LIKHFSGALILIDKMLETTYKIAMDVNTN"  
 1569..2504  
 /gene="a111722"  
 1569..2504  
 /gene="a111722"  
 /note="ORF ID: a111722  
 probable oxidoreductase"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="BAB73421.1"  
 /db\_xref="GI:17130812"  
 /translation="MKNEKMAENIPPOKRLAIVTSSSGIGYETARVLANKQASVI  
 IAVNLDKGNKALAKIILQONKQADVDKMLDELAVLASVKNFAENFRKRYRLDLINN  
 AGVNIIPYSKTTDGFELQCNHGHGALTGOLLEPLISTEGSRIVAVSSGANNMGI  
 DFDLNLNORSYAKKAKVGDSTLANLYFTTYELDKLKDNGIDTLVTASHCPWTLELO  
 RTAGGIYKYLNGIVAODITWCAFLTLRAALEAGLGAIFYGPNGMEMKGIPIVESH  
 ELSKQALAKLWVSEKLTIDVKEFENKKAQSAK"  
 complement(2701..3462)  
 /gene="a111723"  
 complement(2701..3462)  
 /gene="a111723"  
 /note="ORF ID: a111723  
 probable aldolase"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="BAB73422.1"  
 /db\_xref="GI:17130813"  
 /translation="MSVFSRPLPVFNSEIEERLHRKQRLAAFRLLFGKFPESGICAG  
 HITRADPEFTDFHFWNPLGKYFGHIRVSDLVLRNDGVVYGDALINQAPAAHSQIH  
 EAPDVVAARAHISYKAWSSGLRDLPTODACAFEDHALDDYGVVLESEK  
 RLATPILGNKALITLRNHGILTVGHVDEAARFVYLSERSQAQLAAEAAGPKIIGHE  
 TARLTSQVGRPKAGMFSFQPLDYRVAREBELD"  
 complement(3764..5167)  
 /gene="a111724"  
 complement(3764..5167)  
 /gene="a111724"  
 /note="ORF ID: a111724  
 unknown protein"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="BAB73423.1"  
 /db\_xref="GI:17130814"  
 /translation="MSSTSGVPGLEPDLTHPTTELIQSGSLITASLLVPIVALGIA  
 IALMSFARHNSOOTIFIGEMAIRYAOILLEGLOLALVILILVGGFPLSLSRHY  
 HWEOKIAOVVSQSDGLKEQIAPQVRVYTOEPYTNQVANKIVKNDTOOVSLLS  
 LAGSOIQVLEQSDTVQGRSAVRYADYADYKRVNRLKQDINNFFEIOPNGYSLLN  
 KYVERDRRLRELVNPNYSFAFRLEPGQETSRVYVQOGGVNYSAGQLSNFRL  
 TTIANPAGDAFASGVPKERKADGRSTQYTWIFPDNVVSKPFGVTNTPIRQGIIL  
 PRILLALPALTLMWTLLIYLSPLSTINVAIASSIYFACILTLTVLSFMDAQLMTM  
 ISVLITLTWGLGSTRSASLAITITIGAVVPVGLVLPYSGLTSLVAAGLSVIMLA  
 VRHWYNNMIWQEDSKS"  
 complement(5492..5902)  
 /gene="a111725"  
 complement(5492..5902)  
 /gene="a111725"  
 /note="ORF ID: a111725  
 unknown protein"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="BAB73424.1"  
 /db\_xref="GI:17130815"  
 /translation="WFLKFGSLMIASITLNGIAITTSGRSAIEMAKSNPNYLOVA  
 DQRGPIALDASINGTNFTLVQOGEIGIARRRLPAACSQGLFTKKSFLVANGKLT  
 SEDKTEOEIIFPKPTPDAAAMEIVCRVANNHOKK"  
 6321..6842  
 gene  
 CDS

CDS  
 /gene="alr1726"  
 6321..6842  
 /gene="alr1726"  
 /note="ORF\_ID:alr1726"  
 /codon\_start=1  
 /transl\_table=1  
 /product="transposase"  
 /protein\_id="BAB73425.1"  
 /db\_xref="GI:17130816"  
 /translation="MKSQHLQATTLAGEVSOOYLKEETALNALAELODFIMCPDARE  
 VRKALVAVYGVGLYDEIKIQLDVSRSSTGYKCAVEKCIQDLRLNKKRSYINA  
 KQLOBVASHLOTQDCWELGELFQYLAPEYDVVYSKOSTYDLFSEAGISMKTKVNP  
 KADENAVAEKRD"  
 gene  
 /gene="alr1727"  
 6919..7431  
 /gene="alr1727"  
 6919..7431  
 /note="ORF\_ID:alr1727"  
 /codon\_start=1  
 /transl\_table=1  
 /product="transposase"  
 /protein\_id="BAB73426.1"  
 /db\_xref="GI:17130817"  
 /translation="MGGDLGTYWKGSDQETIAVPVNERDKQTYGAVDYILGELVLK  
 AYDGNKNTINLYLEILLANSFDQRLIFMDGAYHSKELRGFLDSVNSGLPTEQWK  
 IHCYRFPNCPQNPDIEDIMLQAKTVWRFCALIPSPSHLKMIFEMFIRHTTDFDTL  
 QMGTYSKIK"  
 gene  
 /gene="alr1728"  
 8064..9401  
 /gene="alr1728"  
 8064..9401  
 /note="ORF\_ID:alr1728"  
 /hypothetical\_protein=""  
 /codon\_start=1  
 /transl\_table=1  
 /protein\_id="BAB73427.1"  
 /db\_xref="GI:17130818"  
 /translation="MTGKNARHNAFLRLVSGNGCAFSGESRYSITSKEVVGDPSCQ  
 VLDAMVRYMSRRAVAVRPVAVSVDKSFVLLDMSANGTYNGQLVGCQHLHG  
 DRISLDGPOPLFEVAVAPQPTITNOVALPBAKHPKPKDVSFPLPITISGCD  
 LTRKAYLVPGILTYVFAVPAFATVGRITQANQVYIATYIATAAYFIIRLCKPKPMW  
 LVGAALSTLILSLPILNLFIFIREVLPSPVPRIQDVSSLTFLFVGYFGAGLMBR  
 LKAKPLIGVLLALGEPARERVGWEPIDGILGTAASVGTLLLETGLQVYPAAS  
 LQAGTEVGLQVLIARILGLPAGHNAVSGYLGFIQLAAKPRHAKQIILAVYLSAAL  
 HALNNTAGHSNNLLVVVGVSYAFMAAILKARALSPTRSQNPATRFIDPKS"  
 gene  
 /gene="alr1729"  
 complement(9666..10310)  
 /complement(9666..10310)  
 /gene="alr1729"  
 /note="ORF\_ID:alr1729"  
 /unknown\_protein=""  
 /codon\_start=1  
 /transl\_table=1  
 /protein\_id="BAB73428.1"  
 /db\_xref="GI:17130819"  
 /translation="MTHTKNGSDACHSVSQMELDLAALLPEPDAAYPMNPDSESE  
 TYFELERKFAQDVSDSEFTTAAVDFYQKLDHMSGTSPSYVNHHTLSVYVHLOEN  
 LHTAFATGIPQNMALATKAAELFASQSGEDLVECVSVLFTWEADDLITAFRPI  
 AYKRSSEGGVLTISVDVNNQEMSALEVAQLSVAIAYAIROLNDSSES"  
 gene  
 /gene="alr1730"  
 complement(10795..12372)  
 /complement(10795..12372)  
 /gene="alr1730"  
 /complement(10795..12372)  
 /gene="alr1730"  
 /note="ORF\_ID:alr1730"  
 /similar\_to\_Help\_protein=""  
 Alignment Scores:  
 Pred. No.: 3.47e+03 Length: 343550  
 Score: 92.50 Matches: 51  
 Percent Similarity: 42.01% Conservative: 41  
 Best Local Similarity: 23.29% Mismatches: 90  
 Query Match: 8.27% Indels: 37  
 DB: 1 Gaps: 10

US-10-034-500-2 (1-218) x AP003587 (1-343550)  
 Oy 1 AAGUUAUThraAlaSerCythrIlyahgValGUseryrAenTyrlEuValAspTyr 20  
 Db 72599 GCGAAATATACCGAATCT--CCCGAGCGAACCCTTAACCTTCCCTTGATATATACCGC 72543  
 Oy 21 SerGlySerMet-----MetMetIyHsValAlaValArgGluProIyGllEgIuEu 38  
 Db 72542 TCAGGTTTATAGGACGCGGCGACGCTTACATCATGCGCGAAG----- 72501  
 Oy 39 AAlaySgluaIaIleuIySIlleHsnAlaIamEcProIyMetSerTyrlGlnIyGly 58  
 Db 72500 GCGGCTGAATCTGTG-----GTAGATCAACTTGACCAAGATATCTCTCACTGCTC 72447  
 Oy 59 LeuTyrrhrheAlaProTySerValIleleProGlnIySerTrpAsnSerCyVal 78  
 Db 72446 GTTACGATGATGACGATGATACGCTGTTCACCCCAACCTGTAACCTGACCAACTGCG 72387  
 Oy 79 AAGUyCyAlaValaIaenThrIleIySerAhpLeuGluIlePhelArgLeuThrPro 98  
 Db 72386 CTCAAAAGTCCATACGTCAGGTAGGCG-----GTCGATTTCTAAC 72342  
 Oy 99 ValGlyAspGly---IleIyMetHsIgluThrValIleangIlyMetProProGlnAla 117  
 Db 72341 TTATCGGGGGATGGCTTAAGGCGTGCAGATATGCAAGCATCACTGCATCCGCAAAA 72282  
 Oy 118 -----AlaValIleleuLeuThrAspGlyIySaenLeuGlyMetAsn-----Pro 133  
 Db 72281 ATTAATCGTGTGCTGCTGCTGCTGATGTCACCGCAATATGCGATTCAGAACCCAAAG 72222  
 Oy 134 ValGluGluValIySerIleTyrlInThrAsnProAsnValCySPhelIValValSer 153  
 Db 72221 ATATCCACGACGACGTCACCAACCAAGCGAAGGATTTACTCAACTTTGGGT 72162  
 Oy 154 PheAlaAspAhpAlaGluGlyIyValaIleIleAspGlnIleValaIleuAsnSerGly 173  
 Db 72161 TTTGCTCAAGGTTTCATGATGATCTACTAATGGATGCGAAGCTGCTAATGCGAAC 72102  
 Oy 174 -----SerValLeuValAspGly 179  
 Db 72101 TTCACTTCATTCAAAGCATGATGACGACAGAGTTTTCAGCATTCGATGATGAT 72042  
 Oy 180 LeuGlnLeuLeu-----GlnAsnProAlaValCySglnGluPheValAsnSerVal 196  
 Db 72041 CTTAGATCTGTAGTACGACAAACTTGAAAGTACACTAGTACGATGATGATC 71985  
 RESULT 14  
 HSIATIH3  
 LOCUS HSIATIH3 2807 bp mRNA linear PRI 30-MAR-1993  
 DEFINITION H sapiens mRNA for inter-alpha-crypsin inhibitor heavy chain H3.  
 ACCESSION X67055  
 VERSION X67055.1 GI:288562  
 KEYWORDS H3 heavy chain; inter-alpha-crypsin inhibitor heavy chain;  
 proteinase inhibitor.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 REFERENCE Bukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 2807)  
 AUTHORS Bourguignon, J.  
 TITLE Direct Submission  
 JOURNAL Submitted (23-JUN-1992) J. Bourguignon, Institut National de la  
 Sante et de la Recherche Medicale INSERM, Unite 295 Faculte de  
 Med-Pharm de Rouen, Ave de l'Universite BP 97, F-76803 St Etienne  
 Rouvray Cedex, FRANCE  
 2 (bases 1 to 2807)  
 REFERENCES Bourguignon, J., Diarra-Mehrpour, M., Thiberville, L., Boes, F.,  
 Seeboue, R. and Martin, J.P.  
 Human pre-alpha-crypsin inhibitor precursor heavy chain. cDNA and  
 deduced amino-acid sequence  
 Eur. J. Biochem. 212 (3), 771-776 (1993)  
 JOURNAL MEDLINE 93215656

PUBMED 7681778 X14690.  
COMMENT See also Location/Qualifiers  
FEATURES  
SOURCE

1. 2807  
/organism="Homo sapiens"  
/isolate="2"  
/db\_xref="taxon:9606"  
/chromosome="3p"  
/map="p211-p212"  
/clone="5"  
/cell\_type="hepatocyte"  
/tissue\_type="liver"  
/clone\_id="lambda gt11"  
/dev\_stage="adult"  
21. 2678  
/cdoon\_start=1  
/product="inter-alpha-trypsin inhibitor heavy chain H3"  
/protein\_id="CAA47439.1"  
/db\_xref="GI:288563"  
/db\_xref="SWISS-PROT:Q06033"  
/translation="WVALSHLSALQLGLCPRRSPRLIGKSLPEGANGIEVYST  
KINSKVTSPAHNVYTMAYNRADRAKESVDEPKPAFTFTLTIDGTVRPNVK  
EKEVAKQYEKASQKTAGLVKASGRKLEKFTVSVAAGSKVTELTYSLEKHK  
KYEYVYLKVPQKLVHFEIVDIEPQGISMLDAEAFITNDLGSALTSGSKG  
HVSFSPSLDQSRSCPTCDSLINGDFTTYDVNRSPNVQIVNGYFVHFAPQGLPV  
VPKNVAFVIDISGMAGKRLKQTKALRILEDMEEDYLNFIIFSGDVSTKEHLVG  
ATPENTLOARFTVSKMEDKMTINDGLRGISMLKAREHRIPERSTSIYIMLTDG  
DANVGSERPKTIOEVNRVAGIGKSPILVNLGNLNTVFLMALLEHNGPARITYES  
DADLOQGFYEVANPLITGVMEYFENAILDITNTTQHPFDSSEIYVAGRVDEDM  
NSFKADVGHGATNDLTTEVEYDKMEKALQERYITGNVIERLWALTLIBOLEKRA  
KNAGEKEKENTLALALDYLDKHFVTLPLSMVTVPEDEDERALADRGDAEATPV  
SPANSYLTQYOPNPYYVDGDPHFIIOIEKDALCFNIDEAGTVLRILQDAATPV  
LTVNQITGDKRSGSPSKRTKYFGKLGIRNOMDFOYETKEIKTCGTRASTFSL  
DTVTITDGLSMINIRKNMVVSGDGVTFVYVILHGVKHPVHRDPLGPPYVDSHRMS  
AQTGELGQFROPFPKYSIDIRPSGDPKPRDATTLYVKNHQLITVTSQSDKTRKDAISG  
TKVCMFVNHNBSGLIDVHDIVPLNF"

BASE COUNT 723 a 732 c 785 g 567 t  
ORIGIN

Alignment Scores:  
Pred. No.: 11 Length: 2807  
Score: 92.00 Matches: 34  
Percent Similarity: 46.72% Conservative: 30  
Best Local Similarity: 24.82% Mismatches: 45  
Query Match: 8.22% Indels: 28  
DB: 9 Gaps: 7

US-10-034-500-2 (1-218) x HSIATTH3 (1-2807)

QY 16 TyrLeuValAspTyrSerGlySerMetMetLeuHisValAlaValArgGluProLys 35  
Db 867 TTTCGATTCAGCATCGCGGCTTCATGCGTGG-----AAA 905  
QY 36 TLeuValAlaLysGluAlaIleLeuLysIleAsnAlaAlaMetProLysMetSerTyr 55  
Db 906 TTAGAGCAGACAAAGAGCGCTTCTCAGATCTCGAAGATATGAAGAGAGCAACTAT 965  
QY 56 GInGlyLLeuTyrThrPheAlaProTyrSerValIleIleProGInGlySerTyrAsn 75  
Db 966 -----CTGAATTTTCATCTGTTCAAGTGAAGATG-----TCCACATGAAA 1007  
QY 76 SerCyValAlaLysGluCyVala-----ValAsnThrIleLysSerAsp 89  
Db 1008 GAGCACTTGTGTCAGGCGACCGCCGAGAACCTCCAGAGAGCCAGACGTTTGAAGAGC 1067  
QY 90 LeuGluIlePheGlyArgLeuThrProValGlyAspGlyIleLysMetHisGluThrVal 109  
Db 1068 ATGAGAGATTAAGAAG--ATGACCAACATCAATGACGGCTGAGAGGGCATCATGATG 1124  
QY 110 IleAsn-----GInMetProProGInAlaAla-----ValIleLeu 121  
Db 1125 CTGAACAGGCGCGAGAGAGAGACAGAAATCCAGAGAGGACCTTCATGTCATCATG 1184

QY 122 LeuThrAspGlyHisAsnAsnLeuGlyMetAsnProValGluGluValLys 138  
Db 1185 CTGACTGATGGGAGATGCCAATGTTGTGAGAGACAGACCCGAAAAATCCAA 1235

RESULT 15  
AR204653  
LOCUS AR204653 3043 bp DNA linear PAT 20-JUN-2002  
DEFINITION Sequence 16 from patent US 6368792.  
ACCESSION AR204653  
VERSION AR204653.1 GI:21502034  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 3043)  
AUTHORS Billing-Medel, P.A., Cohen, M., Colpitts, T.L., Friedman, P.N.,  
Hayden, M., Klass, M.R., Roberts-Repp, L., Russell, J.C. and  
Stroupe, S.D.  
TITLE Reagents and methods useful for detecting diseases of the  
gastrointestinal tract  
Patent: US 6368792-A 16 09-APR-2002;  
JOURNAL Location/Qualifiers  
FEATURES  
SOURCE 1. 3043  
BASE COUNT 1023 a 575 c 608 g 837 t  
ORIGIN

Alignment Scores:  
Pred. No.: 13.7 Length: 3043  
Score: 91.50 Matches: 46  
Percent Similarity: 35.64% Conservative: 26  
Best Local Similarity: 22.77% Mismatches: 55  
Query Match: 8.18% Indels: 75  
DB: 6 Gaps: 8

US-10-034-500-2 (1-218) x AR204653 (1-3043)

QY 17 LeuValAspTyrSerGlySerMetMetLeuHisValAlaValArgGluProLysIle 36  
Db 941 GTTCTGATTAAGTCTGGAAGCATGGG----- 967  
QY 37 GLeuValAlaLysGluAlaIleLeuLysIleAsnAlaAlaMetProLysMetSerTyrGln 56  
Db 968 -----GGTAAAGAGCGCGCTTAATCAATGAATCAAGCAACAAATTTCTGCTGCG 1021  
QY 57 GlyGlyLeuTyrThrPheAlaProTyrSerValIleIleProGInGlySerTyrAsnSer 76  
Db 1022 -----ACTGTTGAAAATGATCTGCGTGGG 1048  
QY 77 CyVal-----AlaGlu 80  
Db 1049 ATGGTTCACTTGATAGTACTGCCACTATTGTAAATTAAGCTAAATCAATAAAGACAGT 1108  
QY 81 CyAlaValAsnThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGly 100  
Db 1109 GATGAAGAAACACACTCTGACAGATTAACCTACATAT-----CCTCGGGA 1156  
QY 101 -----AspGlyIleLysMetHisGluThrValIleAsnGInMetProPro 115  
Db 1157 GGAATCTCATCTGCTGCTGGAATTAATATGATTCATTCAGGTATTTGGAGAGCTTACTTC 1216  
QY 116 GIn-----AlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsn 132  
Db 1217 CAAGTCAGTATCGCAATATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1276  
QY 133 ProValGluGluValIysSerIleTyrGlnThrAsnProAsnValCySerPheHisValVal 152  
Db 1277 TGTATGATGAAGTGA--CAAGTGGGCGCATTTGATTTATTTGCTTGG 1327  
QY 153 SerPheAla----- 155  
Db 1328 GGAAGAGCTGTGATGAGACAGTAATAGAGATGAGACAGATTAACAGAGAAATGCTATTTT 1387

```

Oy 156 -----AspAspAlaGluGlyValAlaIleAepGlnIleValAlaLeuAsnSer 172
Db 1388 TATGTTTCAGATGAAGCTCAGAACATGCGCTTCATGATGCTTTGGGCTCTTACATCA 1447
Oy 173 GlySer 174
Db 1448 GGAAAT 1453

RESULT 16
AR204655
LOCUS DEFINITION Sequence 18 from patent US 6368792.
ACCESSION AR204655
VERSION AR204655.1 GI:21502037
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 3181)
AUTHORS
Bilting-Medel,P.A., Cohen,M., Colpites,T.L., Friedman,P.N.,
Hayden,M., Klags,M.R., Roberts-Rapp,L., Russell,J.C. and
Strope,S.D.
Reagents and methods useful for detecting diseases of the
gastrointestinal tract
Patent: US 6368792-A 18 09-APR-2002;
Location/Qualifiers
1..3181
/organism="unknown"
BASE COUNT 1080 a 593 c 631 g 877 t
ORIGIN

Alignment Scores:
Pred. No.: 14.5 Length: 3181
Score: 91.50 Matches: 46
Percent Similarity: 35.64% Conservative: 26
Best Local Similarity: 22.77% Mismatches: 55
Query Match: 8.18% Indels: 75
Gaps: 8
DB:

US-10-034-500-2 (1-218) x AR204655 (1-3181)
Oy 17 LeuValaApTySerGlySerMetMetClyshValAlaValArgGluProLysIle 36
Db 952 GTTCTTGATAGCTGGAAGCATGGCG----- 978
Oy 37 GluLeuAlaLysGluAlaIleLeuLysIleAenAlaAlaMetProLysMetSerTyGln 56
Db 979 -----GGTAAGACCGCCTTAATGAAATGAAAGCAAGCAAAACATTCTGCTGCAG 1032
Oy 57 GlyGlyLeuTyThrPheAlaProTySerValIleIleProGlnGlySerTyPAsnSer 76
Db 1033 -----ACGTGTAAGAAATGAAATGCTGCGGTGGG 1059
Oy 77 CysVal-----AlaGlu 80
Db 1060 ATGGTCACTTGTAGTAGTACGCACTATTGTAATTAAGCTAATCAATAAAGACAGT 1119
Oy 81 CysAlaValAenThrIleLysSerAspLeuGlnIlePheGlyArgLeuThrProValGly 100
Db 1120 GATGAAAGAAACACACTGATGCGAGATTAACCTACATAT-----CCTTGGA 1167
Oy 101 -----AspGlyIleLysMetHisGluThrValIleAenGlnMetProPro 115
Db 1168 GGAACCTTCATCTGCTGCGAATTAATGATTCAGTGATTTGGAGAGCTACATTCC 1227
Oy 116 Gln-----AlaAlaValIleLeuLeuThrAspGlyHisAenAsnLeuGlyMetAsn 132
Db 1228 CAACTCGATGATCCGAGATGACTGCTGAGTGGGAGATTAACCTCAAGTTCT 1287
Oy 133 ProValGlnGluValLysSerIleTyGlnThrAsnProAsnValCysPheHisValVal 152
Db 1288 TGTATTGATGAAGTAAA-----CAAAAGTGGGCGCATTTTCATTATTTGCTTGG 1338

```

```

Oy 153 SerPheAla----- 155
Db 1339 GGAAGAGCTGCTGATGAAGCAGATTAATAGATGACGAAGATTAACAGAGAGATCATTTT 1398
Oy 156 -----AspAspAlaGluGlyValAlaIleAepGlnIleValAlaLeuAsnSer 172
Db 1399 TATGTTTCAGATGAAGCTCAGAACATGCGCTTCATGATGCTTTGGGCTCTTACATCA 1458
Oy 173 GlySer 174
Db 1459 GGAAAT 1464

RESULT 17
AX322747
LOCUS DEFINITION Sequence 22 from Patent WO0192528.
ACCESSION AX322747
VERSION AX322747.1 GI:18093738
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1
AUTHORS
Macina,R.A., Chen,S.Y., Pluta,J., Sun,Y. and Recipon,H.
TITLE
Method of diagnosing, monitoring, staging, imaging and treating
colon cancer
Patent: WO 0192528-A 22 06-DEC-2001;
Location/Qualifiers
1..3195
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1088 a 596 c 632 g 879 t
ORIGIN

Alignment Scores:
Pred. No.: 14.5 Length: 3195
Score: 91.50 Matches: 46
Percent Similarity: 35.64% Conservative: 26
Best Local Similarity: 22.77% Mismatches: 55
Query Match: 8.18% Indels: 75
Gaps: 8
DB:

US-10-034-500-2 (1-218) x AX322747 (1-3195)
Oy 17 LeuValaApTySerGlySerMetMetClyshValAlaValArgGluProLysIle 36
Db 953 GTTCTTGATAGCTGGAAGCATGGCG----- 979
Oy 37 GluLeuAlaLysGluAlaIleLeuLysIleAenAlaAlaMetProLysMetSerTyGln 56
Db 980 -----GGTAAGACCGCCTTAATGAAATGAAAGCAAGCAAAACATTCTGCTGCAG 1033
Oy 57 GlyGlyLeuTyThrPheAlaProTySerValIleIleProGlnGlySerTyPAsnSer 76
Db 1034 -----ACGTGTAAGAAATGAAATGCTGCGGTGGG 1060
Oy 77 CysVal-----AlaGlu 80
Db 1061 ATGGTCACTTGTAGTAGTACGCACTATTGTAATTAAGCTAATCAATAAAGACAGT 1120
Oy 81 CysAlaValAenThrIleLysSerAspLeuGlnIlePheGlyArgLeuThrProValGly 100
Db 1121 GATGAAAGAAACACACTGATGCGAGATTAACCTACATAT-----CCTTGGA 1168
Oy 101 -----AspGlyIleLysMetHisGluThrValIleAenGlnMetProPro 115
Db 1169 GGAACCTTCATCTGCTGCGAATTAATGATTCAGTGATTTGGAGAGCTACATTCC 1228
Oy 116 Gln-----AlaAlaValIleLeuLeuThrAspGlyHisAenAsnLeuGlyMetAsn 132
Db 1229 CAACTCGATGATCCGAGATGACTGCTGAGTGGGAGATTAACCTCAAGTTCT 1288

```

```

Oy 133 ProValGluGluValIlyseSerIleTyrglnThrAsnProAsnValCysPheHisValVal 152
Db 1289 TGTATTGATGAAGTAAA-----CAAAGTGGGGCCATTGTTTATTATGCTTTG 1339
Oy 153 SerPheAla----- 155
Db 1340 GGAAGAGCTGCTGATGAAGCAATAATAGATGACGAATACAGAGAGAACTATTTT 1399
Oy 156 -----AspAspAlaGluGluIyLysAlaIleIleAspGlnIleValAlaLeuAsnSer 172
Db 1400 TATGTTTCAGATGAAGCTCAGAACAAATGCGCTTCATTGATGCTTTGGGGCTCTTACATCA 1459
Oy 173 Glycer 174
Db 1460 GGAAAT 1465

RESULT 18
AF127035 3204 bp mRNA linear PRI 11-AUG-1999
LOCUS Homo sapiens calcium-activated chloride channel protein 2 (caccc2)
DEFINITION mRNA, complete cds.
ACCESSION AF127035
VERSION AF127035.1 GI:5726288
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 3204)
AUTHORS Agnel, M., Vernat, T. and Culouscou, J.-M.
TITLE Identification of three novel members of the calcium-dependent
chloride channel (CaCC) family predominantly expressed in the
digestive tract and trachea
JOURNAL PEBB Lett. 435 (3), 295-301 (1999)
MEDLINE 99364503
PUBMED 10437792
REFERENCE 2 (bases 1 to 3204)
AUTHORS Agnel, M. and Culouscou, J.-M.
TITLE Direct Submission
JOURNAL Submitted (09-FEB-1999) Groupe Genomique, SYNTHELABO, 10 Rue des
Carrières, Rueil-Malmaison 92500, France
FEATURES
source location/Qualifiers
1..3204
/organism="Homo sapiens"
/db_xref="taxon:9606"
/issue_type="colon"
1..3204
/gene="CaCC2"
29..2782
/gene="CaCC2"
/feature="bovine epithelial chloride channel homolog"
/codon_start=1
/product="calcium-activated chloride channel protein 2"
/protein_id="AAD48398.1"
/db_xref="GI:5726289"
/translation="MGIFRGFVFLVLCILHOSNTSFKLNNNGFEDIVIVDPSVE
DEKITEQIEDMTTASTYIFEXTEKRFPRFANVSILIPKKNENPOYKPKKHNHGHAD
VIVAPPILGRDEPYTKQTECGEGEYIHFPTDLGKQNEVDPGKLFVHEWAL
RWGVDVENEDQPTFRASKKILEATRCASGIGSRKRVYCCGGSCLSRACRIDSTTKL
YKDCQFEPDKQTEKASIMFQSIDSVAFENCKTHNGEASLONIKCNFSTWEVI
SNSEDFKNTIPMTVPPPPVFLIKIRQIVCLVIDKSGMGKDLNNMNOAAHFLV
LQTVNGSVKVMVHPDSTPATVYKLIQKSSDERMTMAGLPTVPLGTSICSGIKYA
FVIGIEHSQDLGSEVLITDDEDNATSSCIDBYKOSGAIYFIALGRAADAVIEMS
KITGSHFYVSDAQNGLIDAFALTSNTDLSQSLQESKGLTNSNANMNTVVI
IDSTYKDFELITWNSLPSSISLMDPSGTEINENFVDAISKAAVLSIPGAKVETVA
VNIQAKANPETLITVTSRAANSVYPIITVNAKMNVDNSFSPMIVYAEIIQGVVPV
ICANVTAFIESONGHTEVELLDNGAGADSPKMDGVSYSTFATYENGSRISKVAVG
GANTARIKIRPLINRAAYIRGVVNGEITANPPEPIDEDTCTTEDEPRTASGAAV
VSGVPSLPLPDQPPSQITDDATYHEDKILITWIRPQNPFGVKYQRIITISILV
DIRDSFDALQVNTDLSPEANSKESFPAKFNENISEENATHIFAIKSIDSNLTSK
VSNIAQVTLFIQANPDIDIPPTPTPTPKSHNSGVNISTLVLSVTSVIVNIFLS
TLI"

```

```

BASE COUNT 1098 a 594 c 633 g 879 t
ORIGIN
Alignment Scores:
Pred. No.: 14.6 Length: 3204
Score: 91.50 Matches: 46
Percent Similarity: 35.64% Conserves: 26
Best Local Similarity: 22.77% Mismatches: 55
Query Match: 8.18% Indels: 75
DB: Gaps: 8
US-10-034-500-2 (1-218) x AF127035 (1-3204)
Oy 17 LeuValaAspTyrSerGlySerMetMetLysHisValAlaValArgLupProLysIle 36
Db 956 GTTCTGATAGTCTGGAAGCATGGG----- 982
Oy 37 GluLeuAlaLysGluAlaIleLeuLysIleAsnAlaAlaMetProLysMetSerTyrgln 56
Db 983 -----GGTAAGAACCGGCTTAATGCAATGATCAACAGCAAAACATTCTCGCTGCAG 1036
Oy 57 GlyGlyLeuTyrrThrPheAlaProTyrSerValIleIleIleProGlnGlySerThrAsnSer 76
Db 1037 -----ACTGTTGAAATGATGATCCTGGGTGGG 1063
Oy 77 CysVal-----AlaGlu 80
Db 1064 ATGCTTCACTTGTGATAGTACTGCCACTATGTGAATAGCTAATCCAAATAAAGCAGT 1123
Oy 81 CysAlaValaAsnThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGly 100
Db 1124 GATGAAGAAACACACTCATGTCAGATGATCTTACCTACATAT-----CCTCTGGGA 1171
Oy 101 -----AspGlyIleLysMetHisGluThrValIleAsnGlnMetProPro 115
Db 1172 GGAATCTCCATCTGCTCTGGAATTAATATGCAATTCACGGATATGAGAGCTTACATTC 1231
Oy 116 Gln-----AlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsn 132
Db 1232 CAATCGATGATGCCGAAGTACTGCTGCTGACTGATGGGAGGATACACATCGCAATGCT 1291
Oy 133 ProValGluGluValIlyseSerIleTyrglnThrAsnProAsnValCysPheHisValVal 152
Db 1292 TGTATTGATGAAGTAAA-----CAAAGTGGGGCCATTGATGATTTATTCCTTTG 1342
Oy 153 SerPheAla----- 155
Db 1343 GGAAGAGCTGCTGATGAAGCAATAATAGATGACGAATACAGAGAACTATTTT 1402
Oy 156 -----AspAspAlaGluGluIyLysAlaIleIleAspGlnIleValAlaLeuAsnSer 172
Db 1403 TATGTTTCAGATGAAGCTCAGAACAAATGCGCTTCATTGATGCTTTGGGGCTCTTACATCA 1462
Oy 173 Glycer 174
Db 1463 GGAAAT 1468

RESULT 19
AK000072 3221 bp mRNA linear PRI 22-FEB-2000
LOCUS Homo sapiens cDNA FLJ20065 f1s, clone COL01613, highly similar to
DEFINITION ECFC BOVIN EPITHELIAL CHLORIDE CHANNEL PROTEIN.
ACCESSION AK000072
VERSION AK000072.1 GI:7019922
KEYWORDS oligo capping; f1s (full insert sequence).
SOURCE Homo sapiens colon cDNA to mRNA, clone_11b.COL clone:COL01613.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (sites)
AUTHORS Kawabata, A., Hiki, J. T., Kobatake, N., Inagaki, H., Ikema, Y.,
Okamoto, S., Okitani, R., Ota, T., Suzuki, Y., Obayashi, M., Nishi, T.,
Shibahara, T., Tanaka, T., Nakamura, Y., Isegai, T. and Sugano, S.

```



TITLE	NEDO human cDNA sequencing project
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 3221)
AUTHORS	Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T., Shihahara,T., Tanaka,T. and Nakamura,Y.
TITLE	Journal
JOURNAL	Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@ime.u-tokyo.ac.jp), Tel:81-3-5449-5266, Fax:81-3-5449-5416)
COMMENT	NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5' - & 3'-end one pass sequencing; Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).
FEATURES	
source	location/Qualifiers 1..3221 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="COL01613" /tissue_type="colon" /clone_id="COL" /note="Cloning vector pMEB8FL3"
misc_feature	1..3221 /note="Highly similar to ECLC_BOVIN EPITHELIAL CHLORIDE CHANNEL PROTEIN"
BASE COUNT	1105 a 600 c 634 g 882 t
ORIGIN	
Alignment Scores:	
Pred. No.:	14.7 Length: 3221
Score:	91.50 Matches: 46
Percent Similarity:	35.64% Conservative: 26
Best Local Similarity:	22.77% Mismatches: 55
Query Match:	8.18% Indels: .75
DB:	Gaps: 9 Gaps: 8
US-10-034-500-2 (1-218) x AKO00072 (1-3221)	
Oy	17 LeuValAapTySerGlySerMetMeCtyBHsValAlaValArgGluProIysile 36
Db	970 GTTCTTGATTAAGTCGTGGAAAGCATGGG----- 996
Oy	37 GluLeuAlaValGluAlaIleLeuIysIleAsnAlaIleMetProIysMetSerTyrgln 56
Db	997 -----GGTAAGACCGCCCTTAATGAATGAATCAAGCAAACAATTCTCGTGCAG 1050
Oy	57 GlyGlyLeuTyThrPheIleProTyrservalIleIleProGlngIyserThrPanser 76
Db	1051 -----ACTGTTGAAAATGATCTGGGTGGG 1077
Oy	77 CysVal-----AlaGlu 30
Db	1078 ATGTTCACTTGATAGTAGTACGTCACATAATTGAAATAGCTAATCCAATAAAGAAGACT 1137
Oy	81 CysAlaValenThrTilleySerAspleuGlulIephnegIyArgleuThyrProValgly 100
Db	1138 GATGAAGAAACACACTCATGTCGACGAGATTACCTACATAT-----CTCTGGGA 1185
Oy	101 -----AspGlyTlleIysMethIsgluThrValIleasnGlnmetPropo 115
Db	1186 GGAAGTTCATTCGCTCGGAATTAAATATGCAATTCAGGTATATGGAAGACTACATTCC 1245
Oy	116 Gln-----AlaAlaValIleleuLeuThraSpGlyHisAsnAsnIleuGlyMetAsn 132
Db	1246 CAATCGATGATCGGAAGTACTGCTGCTGATGATGGGGAAGATPAACACTGCAAGTTCT 1305
Oy	133 ProValGluGluValIylserIleTyrgInThrAsnProAsnValCyserPheHisValVal 152
Db	1306 TGTATTATGAAGTAAA-----CAAGTGGGCCCATTTGATTTATTTGCTTG 1356

[illegible]



```

Db 1228 CAACTCGATGATCCGAATGCTGCTGACTGATGGGAGATACATGCAAGTTCT 1287
Qy 133 ProvalGluGluVallyserIleTyrglnThrAsnProAsnValCySpheniValVal 152
Db 1288 TGTATGATGAAGTGA---CAAGTGGGCCATTGTCATTATTATTCCTTGG 1338
Qy 153 SerPheAla----- 155
Db 1339 GGAAGAGCTGCTGATGAACAGTAATAGATGACGAATACAGAGAACTATTTT 1398
Qy 156 -----AspAspAlaGluGlyLysAlaIleIleAspGlnIleValAlaLeuAsnSer 172
Db 1399 TATGTTTCAGATGAAGCTCAGAAACATGCGCTCATGATGCTTTGGGCTTACATCA 1458
Qy 173 GlySer 174
Db 1459 GGAAAT 1464

RESULT 21
AX376190 3265 bp DNA linear PAT 01-MAR-2002
DEFINITION Sequence 257 from Patent WO0168848.
ACCESSION AX376190
VERSION AX376190.1 GI:19170487
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Baker,K.P., Chen,J., Desnoyers,L., Goddard,A., Godowski,P.J.,
Gurney,A.L., Pan,J., Smith,V., Watanabe,C.K., Wood,W.I. and
Zhang,Z.
Secreted and transmembrane polypeptides and nucleic acids encoding
the same
Patent: WO 0168848-A 257 20-SEP-2001;
JOURNAL
Genentech, Inc. (US)
FEATURES
source location/Qualifiers
1..3265
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1159 a 596 c 632 g 878 t
ORIGIN
Alignment Scores:
Pred. No.: 14.9 Length: 3265
Score: 91.50 Matches: 46
Percent Similarity: 35.64% Conservative: 26
Best Local Similarity: 22.77% Mismatches: 55
Query Match: 8.18% Indels: 75
Gaps: 8
DB: 6

US-10-034-500-2 (1-218) x AX376190 (1-3265)
Qy 17 LeuValAspTyrSerGlySerMetMetLysHisValAlaValArgGluProLysIle 36
Db 952 GTTCTTGATAGCTGGAACATGGG----- 978
Qy 37 GluLeuAlaLysGluAlaIleLeuLysIleAsnAlaIleMetProLysMetSerTyrgln 56
Db 979 -----GTRAAAGACCGCTTAATCGAATCAACAGCAAAACATTTCCGCTGACAG 1032
Qy 57 GlyGlyLeuTyrThrPheAlaProTyrSerValIleIleProGlnGlySerTyrPasnSer 76
Db 1033 -----ACTGTTGAAATGATCTCGGTGGG 1059
Qy 77 CyVal-----AlaGlu 80
Db 1060 ATGCTTCACCTTGATGATGCTGCCACATATTGTAATAGCTAATCCAAATPAAAGCAGT 1119
Qy 81 CyAlaValAsnThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGly 100

```

```

Db 1120 GATGAAGAAACACACTCATGCGAGATTACTTACATAT-----CTCTGGGA 1167
Qy 101 -----AspGlyIleLysMetHisGluThrValIleAsnGlnMetProPro 115
Db 1168 GGAATCTTCATCTGCTCTGGAATTAATATGCTATTCAGGAGATTTGGAGACTTACATCC 1227
Qy 116 Gln-----AlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsn 132
Db 1228 CAACTCGATGATCCGAATGCTGCTGACTGATGGGAGATACATGCAAGTTCT 1287
Qy 133 ProvalGluGluVallyserIleTyrglnThrAsnProAsnValCySpheniValVal 152
Db 1288 TGTATGATGAAGTGA---CAAGTGGGCCATTGTCATTATTATTCCTTGG 1338
Qy 153 SerPheAla----- 155
Db 1339 GGAAGAGCTGCTGATGAACAGTAATAGATGACGAATACAGAGAACTATTTT 1398
Qy 156 -----AspAspAlaGluGlyLysAlaIleIleAspGlnIleValAlaLeuAsnSer 172
Db 1399 TATGTTTCAGATGAAGCTCAGAAACATGCGCTCATGATGCTTTGGGCTTACATCA 1458
Qy 173 GlySer 174
Db 1459 GGAAAT 1464

RESULT 22
AX403491 3265 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 378 from Patent WO0073454.
ACCESSION AX403491
VERSION AX403491.1 GI:21436981
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Ashkenazi,A.J., Baker,K.P., Botstein,D., Desnoyers,L., Eaton,D.,
Ferrara,N., Gerber,H., Gertlesen,M., Goddard,A., Godowski,P.,
Grimaldi,C.J., Gurney,A.L., Kijavich,I., Napier,M.A., Pan,J.,
Paoletti,N.F., Roy,M., Stewart,J.A., Tumas,D., Watanabe,C.K.,
Williams,P., Wood,W.I. and Zhang,Z.
Secreted and transmembrane polypeptides and nucleic acids encoding
the same
Patent: WO 0073454-A 378 07-DEC-2000;
JOURNAL
Genentech Inc. (US)
FEATURES
source location/Qualifiers
1..3265
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1159 a 596 c 632 g 878 t
ORIGIN
Alignment Scores:
Pred. No.: 14.9 Length: 3265
Score: 91.50 Matches: 46
Percent Similarity: 35.64% Conservative: 26
Best Local Similarity: 22.77% Mismatches: 55
Query Match: 8.18% Indels: 75
Gaps: 8
DB: 6

US-10-034-500-2 (1-218) x AX403491 (1-3265)
Qy 17 LeuValAspTyrSerGlySerMetMetLysHisValAlaValArgGluProLysIle 36
Db 952 GTTCTTGATAGCTGGAACATGGG----- 978
Qy 37 GluLeuAlaLysGluAlaIleLeuLysIleAsnAlaIleMetProLysMetSerTyrgln 56
Db 979 -----GTRAAAGACCGCTTAATCGAATCAACAGCAAAACATTTCCGCTGACAG 1032
Qy 57 GlyGlyLeuTyrThrPheAlaProTyrSerValIleIleProGlnGlySerTyrPasnSer 76

```





```

Db 104 GGG-----CCCCAGGGCTCC-----118
Qy 76 SerCyValAlaGluCyAlaValaAsnThrIleYsSerAspLeuGlnIlePheGlyArg 95
Db 119 -----AGCCGGTGCCTTGCTGCCTCAGCCCTCTTCTTCTTCTTGCGCAGAGGCC 169
Qy 96 LeuThr-----ProValGlyAspGlyIleYsMetHisGlu 107
Db 170 ACTAGCCTCTTCTGCTGCTGCACCTTGAGATGATCGC-----CCCAGAGGGA 220
Qy 108 ThrValIleAsnGlnMetPro-----ProGlnAlaValIleLeuLeuThrAspGly 125
Db 221 GAGTTCCTGAATTAACCTCCTCAGCCGCCAGATGCTCACACTCAGATCATCT 280
Qy 126 HisAsnAsnLeuGlyMetAsnProValGluGluValYsSerIleTyGlnThrAsnPro 145
Db 281 TCTCAAAACATGATATGACACAGCTGTAGCC-----310
Qy 146 AsnValCySPheHisValaValaSerPheAlaAspAPAlaGluGlyLeuAlaIleIleAsp 165
Db 311 -----CATGTTGTACCAAAAATGAAGCAAGAG-----340
Qy 166 GlnIleValAlaLeuAsn-----SerGlySerValLeuValAspGlyLeuGlnLeu 183
Db 341 CAGCTGTGTGTGCTAAGTCGTGTCATGCGCTCTGCGCAATGAGCATGAGCTGATA 400
Qy 184 GlnAsnProAlaVal-----CysGlnIleuPheVal 193
Db 401 GACAAACACAGCTGTGTGCTGCGCAACGGGCTATACCTTGTCTACTCCAGGTCTCTTC 460
Qy 194 AsnSerValPheCyGlnGlnIleuGlnIleLeuValThrGluGluVal 208
Db 461 AAGGGCCAAAGGCTGCCCTCCTACGTCGCTCCACCCACACACTGTC 505

RESULT 26
AF012910 1637 bp mRNA linear ROD 19-FEB-2000
LOCUS Marmota monax tumor necrosis factor (TNF) mRNA, complete cde.
DEFINITION AF012910
ACCESSION AF012910.1 GI:7001346
VERSION
KEYWORDS
SOURCE Marmota monax.
ORGANISM Marmota monax.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Scuriidae; Sciurinae;
Marmota.
REFERENCE 1 (bases 1 to 1637)
AUTHORS Li,D.H., Havelil,E.A., Brown,C.L. and Cullen,J.M.
TITL Woodchuck lymphotoxin-alpha, -beta and tumor necrosis factor genes:
structure, characterization and biological activity
JOURN Gene 242 (1-2), 295-305 (2000)
MEDLINE 20184748
PUBMED 10721723
REFERENCE 2 (bases 1 to 1637)
AUTHORS Li,D.H.
TITL Direct Submission
JOURN Submitted (08-JUL-1997) Microbiology, Pathology and Parasitology,
College of Veterinary Medical Sciences, North Carolina State
University, 4700 Hillsborough St., Raleigh, NC 27606, USA
FEATURES
source
1. 1637
/organism="Marmota monax"
/db_xref="taxon:9995"
/cell_type="mononuclear cells"
/tissue_type="peripheral blood"
1. 1637
/gene="TNF"
177. 878
/gene="TNF"
/codon_start=1
/product="tumor necrosis factor"
/protein_id="AAF34863.1"

```

```

/db_xref="GI:7001347"
/translation="MSTESMI RDVELAEALPKEMWPOGSSRCLCLSPFLVWAG
TLECLAHGVIGPQREEPFLNNPLSPQAOMLTLRSSQNNNDKPAHVAKNDKEQ
LWLSRRANMLGIMELIDNOLVVPANGLYLVYSOYLPGGQGPSVLTLHTVSRPA
VSYDKNLNLSAIKSPCKESLBGAERKPRMTERTLYLGVEFLQKDRLSAEVNLPSYL
DFASGQVYFGVIAL"
1311
/gene="TNF"
/feature="AU rich element"
1622. 1637
/gene="TNF"

BASE COUNT 389 a 458 c 387 g 403 t
ORIGIN

Alignment Scores:
Pred. No.: 7.24 Length: 1637
Score: 91.00 Matches: 45
Percent Similarity: 37.44% Conservative: 28
Best Local Similarity: 23.08% Mismatches: 52
Query Match: 8.13% Indels: 70
DB: Gaps: 10

US-10-034-500-2 (1-218) x AF012910 (1-1637)
Qy 36 IlegIleuAlaYsGluAlaIleLeuYsIleAsnAlaIleMetProLYsMetSerTyR 55
Db 204 GTGAGAGTGGCCAGAGAG-----GCACTCCCAAGAGGACATGG 242
Qy 56 GlnGlyIleuTyThrPheAlaProTySerValIleIleProGlnGlySerTyPasn 75
Db 243 GGG-----CCCAGGGCTCC-----257
Qy 76 SerCyValAlaGluCyAlaValaAsnThrIleYsSerAspLeuGlnIlePheGlyArg 95
Db 258 -----AGCCGGTGCCTTGCTGCCTCAGCCCTTCTTCTTCTTGCGCAGAGCC 308
Qy 96 LeuThr-----ProValGlyAspGlyIleYsMetHisGlu 107
Db 309 ACTAGCCTCTTCTGCTGCTGCACCTTGAGATGATCGC-----CCCAGAGGGA 359
Qy 108 ThrValIleAsnGlnMetPro-----ProGlnAlaValIleLeuLeuThrAspGly 125
Db 360 GAGTTCCTGAATTAACCTCCTCAGCCGCCAGATGCTCACACTCAGATCATCT 419
Qy 126 HisAsnAsnLeuGlyMetAsnProValGluGluValYsSerIleTyGlnThrAsnPro 145
Db 420 TCTCAAAACATGATATGACACAGCTGTAGCC-----449
Qy 146 AsnValCySPheHisValaValaSerPheAlaAspAPAlaGluGlyLeuAlaIleIleAsp 165
Db 450 -----CATGTTGTACCAAAAATGAAGCAAGAG-----479
Qy 166 GlnIleValAlaLeuAsn-----SerGlySerValLeuValAspGlyLeuGlnLeu 183
Db 480 CAGCTGTGTGTGCTAAGTCGTGTCATGCGCTCTGCGCAATGAGCATGAGCTGATA 539
Qy 184 GlnAsnProAlaVal-----CysGlnIleuPheVal 193
Db 540 GACAAACACAGCTGTGTGCTGCGCAACGGGCTATACCTTGTCTACTCCAGGTCTCTTC 599
Qy 194 AsnSerValPheCyGlnGlnIleuGlnIleLeuValThrGluGluVal 208
Db 600 AAGGGCCAAAGGCTGCCCTCCTACGTCGCTCCACCCACACACTGTC 644

RESULT 27
AF250878/c 180461 bp DNA circular BCT 16-APR-2002
LOCUS AF250878
DEFINITION Salmonella typhi R27 plasmid complete sequence.
ACCESSION AF250878
VERSION AF250878.1 GI:7800243
KEYWORDS HTG.
SOURCE Salmonella typhi.
ORGANISM Salmonella typhi

```

Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; *Salmonella*.

1 (bases 1 to 180461)  
Sherburne, C.K., Lawley, T.D., Gilmour, M.W., Blattner, F.R.,  
Burland, V., Grobeck, E., Rose, D.J. and Taylor, D.E.  
The complete DNA sequence and analysis of R27, a large IncHI  
plasmid from *Salmonella typhi* that is temperature sensitive for  
transfer

JOURNAL  
MEDLINE  
PUBMED  
20280091  
10773089  
2 (bases 1 to 180461)  
Burland, V., Rose, D.J. and Mayhew, G.F.  
Direct Submission  
Submitted (30-MAR-2000) Laboratory of Genetics, University of  
Wisconsin, 445 Henry Mall, Madison, WI 53706, USA  
R27, a large antibiotic resistance plasmid from *Salmonella typhi*  
location/Qualifiers  
1. 180461  
/organism="Salmonella typhi"  
/db\_xref="taxon:601"  
/plasmid="R27"  
/note="antibiotic resistance plasmid"  
complement(51028..52248)  
/organism="Salmonella typhi"  
/db\_xref="taxon:601"  
/plasmid="R27"  
insertion\_seg="R27"  
/note="antibiotic resistance plasmid"  
74086..75294  
/organism="Salmonella typhi"  
/db\_xref="taxon:601"  
/plasmid="R27"  
insertion\_seg="R27"  
/note="antibiotic resistance plasmid"  
complement(81810..83018)  
/organism="Salmonella typhi"  
/db\_xref="taxon:601"  
/plasmid="R27"  
insertion\_seg="R27"  
/note="antibiotic resistance plasmid"  
complement(86560..158891)  
/organism="Salmonella typhi"  
/db\_xref="taxon:601"  
/plasmid="R27"  
insertion\_seg="R27"  
complement(179821..180461.1..346)  
/gene="R0210"  
/evidence=not experimental  
complement(179821..180461.1..346)  
/gene="R0210"  
/note="spans sequence ends; f202: 21.92 pct identical (1  
gap) to 73 residues of an approx. 1664 aa protein  
GENPEPT: gi|1296953, vitellinogenin. [Oncohychnus mykiss]"  
/codon\_start=1  
/evidence=not experimental  
/transl\_table=11  
/product="unknown"  
/protein\_id="AAF70046.1"  
/db\_xref="GI:7800450"  
/translation="MKQIRPTAFCSQMDPSIFDPFEEREKEIITIVESPDIYINDA  
IAAOSIOSEDHSGPELPSVEDVALPGNCPKLIQKLIHDISPAATDYLTLLE  
YSAELGAIYFMAPEYVQIVYVAMSESLNRPHTVITKSLYSTDPEGNTYNNLEA  
ETRWYKSKFELICATHGIDAIERFSEKLSLYETHSYGENAELARHOKEVILH  
DKEMGWIDIFFEODYLQENKLMKMTORDIMDVYSKVFSTINLFKLVVNNKLT  
RDTFGVIVNPFPERQFEWIESEVEIIVGTHTGVVAAIRQMATCRMTFH"  
complement(407..784)  
/gene="R0001"  
/evidence=not experimental  
complement(407..784)  
/gene="R0001"  
/function="unknown"  
/note="f126; 38.10 pct identical (0 gaps) to 42 residues

of an approx. 464 aa protein GENPEPT: gi|1418563;  
vitellinogenin. [Oncohychnus mykiss]"  
/codon\_start=1  
/evidence=not experimental  
/transl\_table=11  
/product="orf; hypothetical protein"  
/protein\_id="AAF69840.1"  
/db\_xref="GI:7800244"  
/translation="MSYLEYVNEVKVSGIVISAMEKEMTSQNVLLIKLNKMTETIN  
GETERPSIOIKVSPSEMYSGCFGTGNGDELAVSGYLVVDTITLREGRHLDYMRVY  
ATSLTAIRPKPQFGQSSFRQ"  
complement(843..1718)  
/gene="R0002"  
/evidence=not experimental  
complement(843..1718)  
/gene="R0002"  
/note="f292; 24.63 pct identical (1 gap) to 134 residues  
of an approx. 712 aa protein GENPEPT: gi|3608154; unknown  
protein. [Arabidopsis thaliana]"  
/codon\_start=1  
/evidence=not experimental  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="AAF69841.1"  
/db\_xref="GI:7800245"  
/translation="MSKSKARSKALLAFADLIPDMQVYNNKLLDLSLAVYSGHNDL  
IVINMEDOPTIELNSKVSMLAKLSAFSTYVHVMQQLVNPIDFEKAYTLLKEA  
PAIPMFRTIADLDKFLNEBPEKYLGNLFDNDYSLASRELKNQOLVAVSEIIE  
KREKLARNRNETKAYHENVDMVAVPAIMEKIDGPELMHTSELSVPSKGV  
DYAKSKIEPLTKREBOCLDQMKADKROATVDEBLAQINNKKTVKAPRSQGLFGS  
VIAMTWFIVNKEI"  
complement(2116..3900)  
/gene="R0003"  
/note="R0003"  
/evidence=not experimental  
complement(2116..3900)  
/gene="R0003"  
/function="putative DNA - replication, repair,  
restriction/modification"  
/note="f595; 28.29 pct identical (52 gaps) to 555 residues  
of an approx. 728 aa protein GENPEPT: gi|2625021; DNA  
helicase II. [Serratia marcescens]"  
/codon\_start=1  
/evidence=not experimental  
/transl\_table=11  
/product="putative DNA helicase"  
/protein\_id="AAF69842.1"  
/db\_xref="GI:7800246"  
/translation="MGKPTDEQVRIENANANNVIAAPSGCKSFMIETVILIROR  
PYAKGVMTFRAATNSLAELKRLSKQDQVAVTFHGFIRMQLDMVWVGKMLI  
SSAORSVIRALKESGAPFRYPDAEFAIDAGREMDTDIISVRAHROQIHLFNTYQAI  
COKHVALDLNLSRFVGVOMYSGKQFALDHLVVDVQDSDSIQYAVISLHPTAGVN  
TSIYVDDQAIYSPRASGVKIPFOEKEQRPNI FYLNTCRCEBILIKVAGALIEKN  
VYIAKDLRSKAGGSKRTHFRSYVMDQIOGILANTLNODPISMAIISRGNAHLDQLE  
SLIEQPVLYRYGKSFMDKETSVDVHLAEPFRSHDVLMKRVLALFESNEBNDOTE  
LSMKGRKVTFGELNIPNESLETRTLHNLPTQETREKVEIERGFNLKMWELSS  
IKMRQKSPSLIRIADPTCKOMAKQNMNIRAAACGPKKDEEYIPDKVLLS  
TLHGSKGLEKNVLIIMSCNADQIPSKSVGQDAIEERRLLFVGTTRAEOQLHVMWYG  
DPSFPLSCADEKLKKAASKRTESEPLTE"  
4082..4954  
/gene="R0004"  
/evidence=not experimental  
4082..4954  
/gene="R0004"  
/function="unknown"  
/note="f291; 25.71 pct identical (3 gaps) to 105 residues  
of an approx. 1792 aa protein GENPEPT: gi|2621801; DNA  
magnesium chelatase subunit [Methanobacterium  
thermoautotrophicum]"  
/codon\_start=1  
/evidence=not experimental  
/transl\_table=11  
/product="orf; hypothetical protein"

```
/protein_id="AAF69843.1"
/db_xref="GI:7800247"
/translation="MKNRKFSVLSALVASALVACAGNAGCPSNASHDPTGHLVYM
ANAKDSVAKNAYSLAGFDVPSFNEETVCGGADNLSLLSGSGVYTGGLSGISIM
SLGSLVSSDAEIVLQNOVYV/PVPPKTLCPNDSEIVYAGAAVYVNHKESQWNLGN
PSKSSALASCKIDRAVINKSTCELTSEKRPDPVMTSMKTSFOAIRPATETETPOLN
LPBASVSVIRVFP1PFGKNESNNPSG1IFRSDSPKFTPGAAASINGKDYLPFSGE
YGGKGFPERKTLKSK"
complement(4982..8158)
/gene="crhN"
/note="R0005"
/evidence=not experimental
complement(4982..8158)
/gene="crhN"
/function="Plasmid functions; conjugal transfer"
/note="f1059; 32.94 pct identical (3 gaps) to 85 residues
of an approx. 712 aa protein GENPERT: gi|3378368; mating
pair stabilization protein. (Sphingomonas
aromaticivorans)"
/codon_start=1
/evidence=not experimental
/transl_table=1
/product="putative mating pair stabilization protein"
/protein_id="AAF69844.1"
/db_xref="GI:7800248"
/translation="MKRKRITISALIASMINTWMAFYVSLINIALTPKIFADALP
DLENNFLLANPNARNRATTSADIVKYNADSGENLSGKISEKYKASSTINLVG
KYGTPNSESVMNSAVSDKSIKRAVLPMSGGTINENTKEGAKLLSPSGNIGI
SNNPRTAGCTSTGCELPSSRQKSDVQPNAGRGDNGRTINDIKRKSOLFPAQST
DGAIVRTLVANAKENPASTIKPNDPMFAGNEIGONAVAGTGMVQNCSTSTSKTIT
THYDVEKYFNCNPKDNFNSCTTRDSVSVYV1SGGNDMSGDCNCDVAVRGRD
NYMNDGVYDNLTLTKFHPDALATAKIINAMDMHRTVLLGTQI PAIDAVYSSNY
PSPOGMEIKHSKWLKDYDITEQVTSVYOEBREVTMARVWVGSGGEYFEVMT
FENKLEBDKHIQEPAGCEDAVOTPNSPCRPRFVMDVGTRELRSEVTKAATPLYKGD
TGYLTWKNLEGYCDPLAKKELCSYDAKNIMDPTRKDCIYNTDDIKNPDASTY
KNDACVLDNOTCABGFDBGTNTCTYEQKTCDRGDVREVSQTNCSVGM1PCS
GGTEGTEGKEBKDFKGVAAVSNVQYGAKEDEPNDPNSCSVFEKAKWCRSRA
FVNGIATKDCCEKPOGAGSLAELIATLASMIRNTWTVNQLVMTGEGSTASMA
NSVEWMTASAKTGGOMANNYTSLSITSVENYAGNLGRTVSSAAAGSGOLAKETMS
FGGLKQMAKEMAKYDILLPOTVRDPVKNVAVTTGSEIVFSAVONFMAVIVCMTY
AVQYTKMLBMLVACDOKEMASHKRQKSCPTDTRNCYKINLIGFKKVKAKITM
CCTSMLSRVITMOQAYPOLGIDVPVSNCGVSGTQIQCLDPKIDLTEWINDAVQVG
VSPDYATFSESSITANLPFKENYQLPSERTKDMAGSENNIKARQENAOAKIKENDV
CSYLPRAICVSGSYVDPITGKEIPKY"
complement(8181..9128)
/gene="crhJ"
/note="R0006"

gene

Alignment Scores:
Pred. No.: 2.26e+03 Length: 180461
Score: 91.00 Matches: 44
Percent Similarity: 40.46% Conservative: 26
Best Local Similarity: 25.43% Mismatches: 77
Query Match: 8.13% Indels: 26
DB: 1 Gaps: 8

US-10-034-500-2 (1-218) x AP250878 (1-180461)
QY 19 AaPTYSerGlySerMetMetClySHsValAaValArgGluProLysAlaIleGluLeu 38
Db 175610 GATATTTTCAGGTTCAATGATGTCGAACCGCTTATTCATGCAATTAAGATGACCTT 175551
QY 39 AAlaYSGLuaAlaIleLeuYSIleAsnAlaAlaMetProLYsMetSerTYrGlnGly 58
Db 175550 GCATTAACTCTGGCAATTGA-----GCTATTTCAAAATGATCTTCTAATGTCG 175500
QY 59 LeuTYrThrPheAla-----ProTYrSerValIleIleProGlnGlySerTrpAsnSer 76
Db 175499 ATTATTCGCTTGTAGACAAAGATTGTAATTC-----AAA 175461
QY 77 CySvAlaIaGluCyAlaValAsnThrIle---LYsSerAspLeuGluIlePheGlyArg 95
Db 175460 ACGTTTGATGAAACCGCGAGAGAAAGAACTTTCAAATTCCTCACTGGCTGCAAGCAAT 175401
```

```
QY 96 LeuThrProValGlyYAspGlyIleLYsMetCHis---GluThrValIleAsnGlnMetPro 114
Db 175400 AATACCCCAAGCGGCTTCGCAATTAAAGCTGCTTTAGAGTTGTTACTTGAAAGCCAGTTT 175341
QY 115 ProGlnAlaAlaValIleLeuLeuThrAspGlyYHisAsnAsnLeuGlyMetAsnProVal 134
Db 175340 GACAGCAAAATGTTGTTCTTATCTATGATGCTTATCCAGTACGATCGGCTTACACCAAT 175281
QY 135 GluGluValLYsSerIleTYrGlnThrAsnProAsnValCySPheHisValValSerPhe 154
Db 175280 AATGATGTTTTCAGTTCGCTGAAAGTAAAC----- 175251
QY 155 AAlaSPaSPaIaGluGlyLYsAlaIle---IleAspGlnIleValAlaLeuAsnSerGly 173
Db 175250 GGCATTGAATACGCGGTGCTGATTTAAACGTAGTGTGATGATGCGTTTAAATAGAGGC 175191
QY 174 Ser---ValLeuValAspGlyLeuGlnLeuGlnAsn 185
Db 175190 ACGTTCGTAAATGTTGATGATATTTCTTATTCGCCCAAT 175152

RESULT 28
STYPRHCM1
LOCUS 218160 bp DNA linear BCT 26-OCT-2001
DEFINITION Salmonella enterica serovar Typhi (Salmonella typhi) strain CT18
ACCESSION AL513383
VERSION AL513383.1 GI:16505740
KEYWORDS
SOURCE
ORGANISM
Salmonella enterica subsp. enterica serovar Typhi.
Salmonella enterica subsp. enterica serovar Typhi
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Salmonella.
REFERENCE
1 (bases 1 to 218160)
Parkhill,J., Dougan,G., James,K.D., Thomson,N.R., Pickard,D.,
Wain,J., Churcher,C., Mungall,K.L., Bentley,S.D., Holden,M.T.G.,
Sebahia,M., Baker,S., Baaham,D., Brooks,K., Chillingworth,T.,
Comerton,P., Cronin,A., Davis,P., Davies,R.M., Dowd,L., White,N.,
Farar,J., Feltwell,T., Hamlin,N., Haque,A., Hien,T.T., Holroyd,S.,
Jagels,K., Krogg,A., Larsen,T.S., Leather,S., Moule,S., O'Georga,P.,
Perry,C., Quail,M., Rutherford,K., Simmonds,M., Skelton,J.,
Stevens,K., Whitehead,S. and Barrrell,B.G.
Complete genome sequence of a multiple drug resistant Salmonella
enterica serovar Typhi CT18
Nature 413 (6858), 848-852 (2001)
11677608
PUBMED
2 (bases 1 to 218160)
Parkhill,J.
Direct Submission
Submitted (25-OCT-2001) Submitted on behalf of the Salmonella
sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
Notes:
Details of S. typhi sequencing at the Sanger Centre are available
on the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/S_typhi/).

FEATURES
SOURCE
1..218160
/organism="Salmonella enterica subsp. enterica serovar
Typhi"
/db_xref="taxon:90370"
/plasmid="rHCM1"
complement(1..528)
/gene="HCM1.01c"
complement(1..528)
/partial
/gene="HCM1.01c"
/note="HCM1.01c, possible membrane protein, len: 185 aa;
unknown function, contains hydrophobic, possible
membrane-spanning regions. Spans sequence end"
/codon_start=1
/transl_table=1
```

```

RBS
/product="putative membrane protein"
/protein_id="CAD09619.1"
/db_xref="GI:16505741"
/translation="MNQETGSPGFETTVDDEERIAIRWIGVVGVRKQVPPDRVQ
KATPEVNGGLMALALMLALMLVLCIIQRLATQDRMSFGEEDVADIQFTMLVIEFAL
GNVWVFILATGEFSRPMVYRLPLPAKPVWLFDVEVDADVLISGNPFIKWLHKELIKI
RLATYTRLEATGESEKIA"
complement(536..540)
/note="possible RBS"
complement(742..1053)
/gene="HCM1.02c"
complement(742..1053)
/note="HCM1.02c, hypothetical protein, len: 103 aa;
unknown function"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="CAD09620.1"
/db_xref="GI:16505742"
/translation="MMYPDITGVKKRLSOLEVGMTVSLSTPKNGNTTIKTVILHSV
IKEVADNHVITSWNGNAPRRPFETAITGMKKKKPLLIRDSGRARLATREKARILDT
K"
complement(1367..1690)
/gene="HCM1.03c"
complement(1367..1690)
/gene="HCM1.03c"
/note="HCM1.03c, hypothetical protein, len: 107 aa;
unknown function"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="CAD09621.1"
/db_xref="GI:16505743"
/translation="MNNKPLVSPAEISGNAINVAROSVIDMEMDAREKIGKARSLFSS
GIRHNVNGPILQSANQLAVIKRLGPTKYLDCITENLCMFPEGLYVIFMGRRTI
NEPVA"
complement(1697..1702)
/note="possible RBS"
complement(1871..2585)
/note="IS1"
1871..1883
/note="13 bp inverted repeat flanking IS1"
complement(1876..2346)
/gene="insB"
/note="HCM1.04c"
complement(1876..2346)
/gene="insB"
/note="HCM1.04c"
complement(1876..2346)
/note="insB"
/note="HCM1.04c, insB, possible IS1 transposase, len: 156
aa; highly similar to many from Enterobacteriaceae e.g.
SW:IS1_ECOLI (EMBL:X52534), insA, Escherichia coli
insertion element IS1 protein insA (91 aa), fastr scores;
E(): 0, 98.9% identity in 91 aa overlap. Contains PS00017
ATP/GTP-binding site motif A (P-loop). Contains probable
helic-turn-helix motif at aa 67-88"

CDS
complement(2265..2540)
/gene="insA"
/note="HCM1.05c, insA, probable IS1 transposase, len: 91
aa; highly similar to many from Enterobacteriaceae e.g.
SW:IS1_ECOLI (EMBL:X52534), insA, Escherichia coli
insertion element IS1 protein insA (91 aa), fastr scores;
E(): 0, 98.9% identity in 91 aa overlap. Contains PS00017
ATP/GTP-binding site motif A (P-loop). Contains probable
helic-turn-helix motif at aa 67-88"

gene
complement(2265..2540)
/gene="insA"
/note="HCM1.05c"
complement(2265..2540)
/gene="insA"
/note="HCM1.05c, insA, probable IS1 transposase, len: 91
aa; highly similar to many from Enterobacteriaceae e.g.
SW:IS1_ECOLI (EMBL:X52534), insA, Escherichia coli
insertion element IS1 protein insA (91 aa), fastr scores;
E(): 0, 98.9% identity in 91 aa overlap. Contains PS00017
ATP/GTP-binding site motif A (P-loop). Contains probable
helic-turn-helix motif at aa 67-88"

CDS
complement(2265..2540)
/gene="insA"
/note="HCM1.05c"
complement(2265..2540)
/gene="insA"
/note="HCM1.05c, insA, probable IS1 transposase, len: 91
aa; highly similar to many from Enterobacteriaceae e.g.
SW:IS1_ECOLI (EMBL:X52534), insA, Escherichia coli
insertion element IS1 protein insA (91 aa), fastr scores;
E(): 0, 98.9% identity in 91 aa overlap. Contains PS00017
ATP/GTP-binding site motif A (P-loop). Contains probable
helic-turn-helix motif at aa 67-88"

RBS
/product="putative membrane protein"
/protein_id="CAD09619.1"
/db_xref="GI:16505741"
/translation="MNQETGSPGFETTVDDEERIAIRWIGVVGVRKQVPPDRVQ
KATPEVNGGLMALALMLALMLVLCIIQRLATQDRMSFGEEDVADIQFTMLVIEFAL
GNVWVFILATGEFSRPMVYRLPLPAKPVWLFDVEVDADVLISGNPFIKWLHKELIKI
RLATYTRLEATGESEKIA"
complement(536..540)
/note="possible RBS"
complement(742..1053)
/gene="HCM1.02c"
complement(742..1053)
/note="HCM1.02c, hypothetical protein, len: 103 aa;
unknown function"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="CAD09620.1"
/db_xref="GI:16505742"
/translation="MMYPDITGVKKRLSOLEVGMTVSLSTPKNGNTTIKTVILHSV
IKEVADNHVITSWNGNAPRRPFETAITGMKKKKPLLIRDSGRARLATREKARILDT
K"
complement(1367..1690)
/gene="HCM1.03c"
complement(1367..1690)
/gene="HCM1.03c"
/note="HCM1.03c, hypothetical protein, len: 107 aa;
unknown function"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="CAD09621.1"
/db_xref="GI:16505743"
/translation="MNNKPLVSPAEISGNAINVAROSVIDMEMDAREKIGKARSLFSS
GIRHNVNGPILQSANQLAVIKRLGPTKYLDCITENLCMFPEGLYVIFMGRRTI
NEPVA"
complement(1697..1702)
/note="possible RBS"
complement(1871..2585)
/note="IS1"
1871..1883
/note="13 bp inverted repeat flanking IS1"
complement(1876..2346)
/gene="insB"
/note="HCM1.04c"
complement(1876..2346)
/gene="insB"
/note="HCM1.04c"
complement(1876..2346)
/gene="insB"
/note="HCM1.04c, insB, possible IS1 transposase, len: 156
aa; highly similar to many from Enterobacteriaceae e.g.
SW:IS1_ECOLI (EMBL:DI0483), insB, Escherichia coli
insertion element IS1 protein insB (167 aa), fastr scores;
E(): 0, 91.6% identity in 167 aa overlap"
/codon_start=1
/transl_table=1
/product="putative IS1 transposase"
/protein_id="CAD09622.1"
/db_xref="GI:16505744"
/translation="MNGNRPHYGRMDPQDFTSFKLLRPQSVYTRTIQDGPDIYVAKSKS
QKRLPAYARLAKTYVAHYAFGERGMAITGRILMSLSPDVYIMTDCPIFYESLKLKGG
LHVISKRYTORLERHNLNRQRLARLGRKSLSFSSVBEUHDVYIGHYLNITHYO"
complement(2265..2540)
/gene="insA"
/note="HCM1.05c"
complement(2265..2540)
/gene="insA"
/note="HCM1.05c, insA, probable IS1 transposase, len: 91
aa; highly similar to many from Enterobacteriaceae e.g.
SW:IS1_ECOLI (EMBL:X52534), insA, Escherichia coli
insertion element IS1 protein insA (91 aa), fastr scores;
E(): 0, 98.9% identity in 91 aa overlap. Contains PS00017
ATP/GTP-binding site motif A (P-loop). Contains probable
helic-turn-helix motif at aa 67-88"

CDS
complement(2265..2540)
/gene="insA"
/note="HCM1.05c"
complement(2265..2540)
/gene="insA"
/note="HCM1.05c, insA, probable IS1 transposase, len: 91
aa; highly similar to many from Enterobacteriaceae e.g.
SW:IS1_ECOLI (EMBL:X52534), insA, Escherichia coli
insertion element IS1 protein insA (91 aa), fastr scores;
E(): 0, 98.9% identity in 91 aa overlap. Contains PS00017
ATP/GTP-binding site motif A (P-loop). Contains probable
helic-turn-helix motif at aa 67-88"

RBS
/product="putative membrane protein"
/protein_id="CAD09619.1"
/db_xref="GI:16505741"
/translation="MNQETGSPGFETTVDDEERIAIRWIGVVGVRKQVPPDRVQ
KATPEVNGGLMALALMLALMLVLCIIQRLATQDRMSFGEEDVADIQFTMLVIEFAL
GNVWVFILATGEFSRPMVYRLPLPAKPVWLFDVEVDADVLISGNPFIKWLHKELIKI
RLATYTRLEATGESEKIA"
complement(536..540)
/note="possible RBS"
complement(742..1053)
/gene="HCM1.02c"
complement(742..1053)
/note="HCM1.02c, hypothetical protein, len: 103 aa;
unknown function"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="CAD09620.1"
/db_xref="GI:16505742"
/translation="MMYPDITGVKKRLSOLEVGMTVSLSTPKNGNTTIKTVILHSV
IKEVADNHVITSWNGNAPRRPFETAITGMKKKKPLLIRDSGRARLATREKARILDT
K"
complement(1367..1690)
/gene="HCM1.03c"
complement(1367..1690)
/gene="HCM1.03c"
/note="HCM1.03c, hypothetical protein, len: 107 aa;
unknown function"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="CAD09621.1"
/db_xref="GI:16505743"
/translation="MNNKPLVSPAEISGNAINVAROSVIDMEMDAREKIGKARSLFSS
GIRHNVNGPILQSANQLAVIKRLGPTKYLDCITENLCMFPEGLYVIFMGRRTI
NEPVA"
complement(1697..1702)
/note="possible RBS"
complement(1871..2585)
/note="IS1"
1871..1883
/note="13 bp inverted repeat flanking IS1"
complement(1876..2346)
/gene="insB"
/note="HCM1.04c"
complement(1876..2346)
/gene="insB"
/note="HCM1.04c"
complement(1876..2346)
/gene="insB"
/note="HCM1.04c, insB, possible IS1 transposase, len: 156
aa; highly similar to many from Enterobacteriaceae e.g.
SW:IS1_ECOLI (EMBL:DI0483), insB, Escherichia coli
insertion element IS1 protein insB (167 aa), fastr scores;
E(): 0, 91.6% identity in 167 aa overlap"
/codon_start=1
/transl_table=1
/product="putative IS1 transposase"
/protein_id="CAD09622.1"
/db_xref="GI:16505744"
/translation="MNGNRPHYGRMDPQDFTSFKLLRPQSVYTRTIQDGPDIYVAKSKS
QKRLPAYARLAKTYVAHYAFGERGMAITGRILMSLSPDVYIMTDCPIFYESLKLKGG
LHVISKRYTORLERHNLNRQRLARLGRKSLSFSSVBEUHDVYIGHYLNITHYO"
complement(2265..2540)
/gene="insA"
/note="HCM1.05c"
complement(2265..2540)
/gene="insA"
/note="HCM1.05c, insA, probable IS1 transposase, len: 91
aa; highly similar to many from Enterobacteriaceae e.g.
SW:IS1_ECOLI (EMBL:X52534), insA, Escherichia coli
insertion element IS1 protein insA (91 aa), fastr scores;
E(): 0, 98.9% identity in 91 aa overlap. Contains PS00017
ATP/GTP-binding site motif A (P-loop). Contains probable
helic-turn-helix motif at aa 67-88"

CDS
complement(2265..2540)
/gene="insA"
/note="HCM1.05c"
complement(2265..2540)
/gene="insA"
/note="HCM1.05c, insA, probable IS1 transposase, len: 91
aa; highly similar to many from Enterobacteriaceae e.g.
SW:IS1_ECOLI (EMBL:X52534), insA, Escherichia coli
insertion element IS1 protein insA (91 aa), fastr scores;
E(): 0, 98.9% identity in 91 aa overlap. Contains PS00017
ATP/GTP-binding site motif A (P-loop). Contains probable
helic-turn-helix motif at aa 67-88"

RBS
/product="putative membrane protein"
/protein_id="
```

```

misc_feature
/codon_start=1
/transl_table=1
/product="putative IS1 transposase"
/protein_id="CAD09623.1"
/db_xref="GI:16505745"
/translation="MASVSISSGSCATDGVNNGKSTAGHQRYLCSHCKRTWQLOFT
YTASPGHOKIIMAMGVGCACATATIMGVGLNTILRLHKNGSRSR"
complement(2284) .2289)
/gene="tnaA"
/notes="possible translational frameshift site, similar to
that determined experimentally (EMBL:X52534)"
complement(2472) .2495)
/gene="tnaA"
/notes="PS00017 ARP/GTP-binding site motif A (P-loop)"
complement(2573) .2585)
/notes="13 bp inverted repeat flanking IS1"
complement(2694) .3644)
/gene="CORa"
/notes="HCM1.07C"
complement(2694) .3644)
/gene="CORa"
/notes="HCM1.07C, corA, probable magnesium and cobalt
transport protein, len: 316 aa; similar to many e.g.
SM:CORa ECOLI (EMBL:L11042), corA, Escherichia coli
magnesium and cobalt transport protein (316 aa), fasta
scores; E(): 0, 99.1% identity in 316 aa overlap. Contains
Pfam match to entry PF01544 CorA, CorA-like Mg2+
transporter protein. Contains hydrophobic, possible
membrane-spanning regions"
/codon_start=1
/transl_table=1
/product="putative magnesium and cobalt transport protein"
/protein_id="CAD09624.1"
/db_xref="GI:16505746"
/translation="MLSAFOLENNRLTRLEVESQPLVNAWIDLYEPDDDERLRQGS
ELGGSILARPELDEIASARFEDDDGLIHSPFFPADAHGNSVATIDGRLEF
TLRREELPAPFLYPMARSOSMDNGMAYELLDLPFKRIQADENITVSDLEQSR
VIMRHQSDYDERLSTLAELBDIGKVRICLMDTQALNPLFKARLPBGQLEQARE
ILRIEQLIPNESLFRQVNFVQAMGFNIEQNRIKLFSVSVVLEPPTLVASSY
GMNPFEMELTMRGYPALIFMLIAGIAPLYFKRKNWP"
complement(2697) .3584)
/gene="CORa"
/notes="Pfam match to entry PF01544 CorA, CorA-like Mg2+
transporter protein, score 432.90, E-value 2.8e-126"
complement(3653) .3656)
/notes="possible RBS"
4004. .4007
/notes="possible RBS"
4014. .4343
/gene="HCM1.08"
4014. .4343
/gene="HCM1.08"
/notes="HCM1.08, hypothetical protein, len: 109 aa; similar
to the N-terminal half of SM:YJGE ECOLI (EMBL:I02122),
yJge, Escherichia coli hypothetical protein (254 aa),
fasta scores; E(): 0, 97.9% identity in 95 aa overlap"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="CAD09625.1"
/db_xref="GI:16505747"
/translation="MAHQLIGKMITLTKRITLALTLPLFVAADDCALSDPTIT
VOAYVAVQTRVVMQKANGEMGTLHLALADINSQGOVAMNGGIINTLVFDIK
YRKREVP"
complement(4363) .4569)
/gene="HCM1.10C"
complement(4363) .4569)
/gene="HCM1.10C"
/notes="HCM1.10C, hypothetical protein, len: 68 aa; unknown
function"
/codon_start=1
/transl_table=1
/product="hypothetical protein"

```



/protein\_id="CAD09626.1"  
 /db\_xref="GI:16505748"  
 /translation="MGHVDYTRTRVOLYDASRPHDGAEOAGELHTVAFSPKPAIAD  
 DIQKIVDTVAEVLGKRYSVNVFSN"

## Alignment Scores:

Pred. No.: 2.85e+03 Length: 218160  
 Score: 91.00 Matches: 44  
 Percent Similarity: 40.46% Conservative: 26  
 Best Local Similarity: 25.43% Mismatches: 77  
 Query Match: 8.13% Indels: 26  
 DB: 1 Gaps: 8

US-10-034-500-2 (1-218) x STYPPHQM (1-218160)

QY 19 AAPTYSerGlySerMetMetLeuHisValAlaValArgGluProLysIleGluLeu 38  
 Db 91039 GATATTTCAGGTTCAATAGCTGAGAACCCGTTATTCATGCAATTAGACTGACCTT 91098  
 QY 39 AlAluysGluAlaIleLeuLysIleAsnAlaAlaMetProLysMetSerTyrGlnGly 58  
 Db 91099 GCATTAACCTCGCAATGAA-----GCTATTCAAAAATGCAATGCTGTAATG 91149  
 QY 59 LeuTyrThrPheAla-----ProTyrSerValIleIleProGlnGlySerTrpAsnSer 76  
 Db 91150 ATTATCCGTTGTAGACAAAGATTGAAATTATC-----AAA 91188  
 QY 77 CyValAlaGluCysAlaValAsnThrIle---LysSerAspLeuGluIlePheGlyArg 95  
 Db 91189 ACGTTTGATGAAAGCGCGAAGAGAACTTCAAAATTCCTCCTGGCTGCAAGGCAAT 91248  
 QY 96 LeuThrProValGlyAspGlyIleLysMetHis---GluThrValIleAsnGlnMetPro 114  
 Db 91249 AATACCCCAAGCGGTTGTCATTAATGCTGCTTAGAGTTGTTACTTGAAGCCAGTTT 91308  
 QY 115 ProGlnAlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsnProVal 134  
 Db 91309 GACAGGAAATTTGTTTCTTATTAAGTATGATGCTTACGACTGAGTCCGCTTACACAT 91368  
 QY 135 GluGluValLysSerIleTyrGlnThrAsnProAsnValCysPheHisValAlaSerPhe 154  
 Db 91369 AATGATGTTTTCAGCTGCTGCAAGTAAAC----- 91398  
 QY 155 AlaAspAspAlaGluGlyLysAlaIle---IleAspGlnIleValAlaLeuAsnSerGly 173  
 Db 91399 GGCATTTGAATCGCTGCTGCTGATTAAGTATGATGCTGCTTAAATGAGGCG 91458  
 QY 174 Ser---ValLeuValAspGlyLeuGlnLeuLeuGlnAsn 185  
 Db 91459 ACCTTCGTAATGTTGATGATATTTCTTATTCGCCAAT 91497  
 RESULT 29  
 LOCUS AF413571 2898 bp DNA linear INV 01-MAY-2002  
 DEFINITION Cryptosporidium parvum cGMP-dependent protein kinase gene, complete  
 ccd.  
 ACCESSION AF413571  
 VERSION AF413571.1 GI:20378273  
 KEYWORDS  
 SOURCE Cryptosporidium parvum.  
 ORGANISM Cryptosporidium parvum.  
 Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;  
 Cryptosporididae; Cryptosporidium.  
 1 (bases 1 to 2898)  
 Gurnett,A.M., Liberator,P.A., Dulski,P.M., Salowe,S.P.,  
 Donald,R.G.K., Anderson,J.W., Wiltse,V., Diaz,C.A., Harris,G.,  
 Chang,B., Darkin-Ratcliff,S.J., Nare,B., Crumley,T., Blum,P.S.,  
 Meunier,A.S., Tames,T., Sardana,M.K., Yuan,J., Biffo,T. and  
 Schmeatz,D.M.  
 Purification and molecular characterization of cGMP-dependent  
 protein kinase from Apicomplexan parasites. A novel  
 chemotherapeutic target  
 J. Biol. Chem. 277 (18), 15913-15922 (2002)

MEDLINE 21975164  
 PUBMED 11834729  
 REFERENCE 2 (bases 1 to 2898)  
 AUTHORS Donald,R.G.K., Liberator,P.A.  
 TITLE Submitted (22-AUG-2001) Human and Animal Infectious Disease  
 Direct Submission  
 JOURNAL Research, Merck & Co., Inc., 126 East Lincoln Avenue, Rahway, NJ  
 07065, USA

## FEATURES

## source

1..2898  
 /organism="Cryptosporidium parvum"  
 /db\_xref="taxon:5807"  
 <1..>2898  
 /product="cGMP-dependent protein kinase"

## mRNA

1..2898  
 /codon\_start=1  
 /product="cGMP-dependent protein kinase"  
 /protein\_id="P4M20902.1"  
 /db\_xref="GI:20378274"

## CDS

/translation="MAFLTRLKTSKFSGGVAAKETTYSLEKEENDEIYEPEKRVV  
 FDRORAKIORSIGLSKMSNISNAGCGMEDYEONVMSHLKORKEDEDIKTSK  
 ALAGNVVAGLNSERATIVLSMRYVEYEGAVI ECGASGFYFVSTSGVGEING  
 NRVTMSGTRAPGELAIHNTPRSATLIVIEKGJMGCGSTPDTLRLLSRVREN  
 RAFLIESLISGLDQKSLSEALVEIEVKQVILIREKIGNVLTMTSGIVGEV  
 EDKYSRLNEDGDAFGERSLMEDEPRASIVANATECLTLNRGILTLGNGVLSK  
 NLIOQSLONSPIFKQPTKQMOILDKISITKQKDVILSTAEKALMAEAFVLEGS  
 IVYRDLKPENILDSOGYIKLIDEGCAKIKLAGRSYTLATGPVMAPEVILGKYNLSC  
 DANAIGICLVIEFGCGLPFGNDVNDHIEFDITSLTSLVPRPHLNDVDVNIIRKLC  
 RPEVSRGSCATGYKKEIKQNVFDPDRPDLRLGRSYAPLVRKTRVAKNTENIYIE  
 NITVEVSNEELSPNTHNITSKQISFNASKNHSVSTRGNMDEYDWIEF"

BASE COUNT 995 a 374 c 609 g 920 t  
 ORIGIN

## Alignment Scores:

Pred. No.: 16.4 Length: 2898  
 Score: 90.50 Matches: 41  
 Percent Similarity: 40.19% Conservative: 43  
 Best Local Similarity: 19.62% Mismatches: 82  
 Query Match: 8.09% Indels: 43  
 DB: 3 Gaps: 7

US-10-034-500-2 (1-218) x AF413571 (1-2898)

QY 19 AAPTYSerGlySerMetMetLeuHisValAlaValArgGluProLysIleGluLeu 37  
 Db 229 GATTTGAAACGAATATGCTAATCTCACTTAAAGATAGAGAAAGACAGAGAGAC 288  
 QY 38 -----LeuAlaLysGlu 41  
 Db 289 ATAAAACTATTCCAAAGCATTTGGACAGAAAGTAAAGGAGCTTCACTGAATGAAATCT 348  
 QY 42 AlAlaIleLeuLysIleAsnAlaAlaMetProLysMetSerTyrGlnGlyLeuTyrThr 61  
 Db 349 GAAATTCCTACTTGGCTGAGTTCATGATTAAGATTAAGATGAGGCG----- 399  
 QY 62 PheAlaProTyrSerValIleIleProGlnGlySerTrpAsnSerCysValAlaGluCys 81  
 Db 400 -----GAAGTGTAAATTAGACAGAGACA-----TTCGATTTTACTTTT 441  
 QY 82 AlAlaValAsnThrIleLysSerAspLeuGluIlePheGly---ArgLeuThrProValGly 100  
 Db 442 GTCAATTTCAACAGAAAGTTTGGGTTGATGATTAAGTAAACAGAGTCAACACAAATGCA 501  
 QY 101 AspGlyIleLysMetHisGluThrValIleAsnGlnMetProProGlnAlaAlaValIle 120  
 Db 502 GAAGGACAGCTTTTGGAGAAATTAGCGCTTATTCACAATACTCCAGAAAGTCAACAT 561







MEDLINE 21301939  
PubMed 11407914  
REFERENCE 3 (bases 1 to 51158)  
AUTHORS Simpson,A.J.G., Reimach,F.C., Arruda,P., Abreu,F.A., Acencio,M., Alvarenga,R., Alves,L.M.C., Araya,J.E., Bala,G.S., Baptista,C.S., Barros,M.H., Bonaccorsi,E.D., Bordin,S., Bove,J.M., Brito,M.R.S., Bueno,M.R.P., Camargo,A.A., Camargo,L.E.A., Carraro,D.M., Carrer,H., Colauto,N.B., Colombo,C., Costa,F.F., Costa,M.C.R., Costa-Neto,C.M., Coutinho,L.L., Cristofani,M., Dias-Neto,E., Docena,C., El-Dorry,H., Facincani,A.P., Ferreira,A.J.S., Ferreira,V.C.A., Ferro,J.A., Fraga,J.S., Franca,S.C., Franco,M.C., Frhmel,M., Gubian,L.R., Garnier,M., Goldman,G.H., Goldman,M.H.S., Gomes,S.L., Gruber,A., Ho,P.L., Hohenisel,J.D., Junqueira,M.L., Kemper,E.L., Kitajima,J.P., Krieger,J.E., Kuramae,E.E., Laigret,F., Lambas,M.R., Leite,L.C.C., Lemos,E.G.M., Lemos,M.V.F., Lopes,S.A., Lopes,C.R., Machado,J.A., Machado,M.A., Madeira,A.M.B.N., Madeira,H.M.F., Marino,C.L., Marques,M.V., Martins,E.A.L., Martins,E.M.F., Matsukuma,A.Y., Menck,C.F.M., Miracca,E.C., Miyaki,C.Y., Monteiro-Vitorello,C.B., Moon,D.H., Nagai,M.A., Nascimento,A.L.T.O., Netto,L.E.S., Nhani Jr.,A., Nobrega,F.G., Nunes,L.R., Oliveira,M.A., de Oliveira,M.C., de Oliveira,R.C., Palmieri,D.A., Paris,A., Peixoto,B.R., Pereira,G.A.G., Pereira Jr.,H.A., Pequeiro,J.B., Quaggio,R.B., Roberto,P.G., Rodrigues,V., de M. Rosa,A.J., de Rosa Jr.,V.E., de Sa,R.G., Santelli,R.V., Sawasaki,H.E., da Silva,A.C.R., da Silva,F.R., da Silva,A.M., Silva Jr.,W.A., da Silveira,J.F., Silveira,M.L.Z., Siqueira,W.J., de Souza,A.A., de Souza,A.P., Terenzi,M.F., Truffi,D., Tsai,S.M., Tsubako,M.H., Vallada,H., van Sluys,M.A., Verjovski-Almeida,S., Vettore,A.L., Zago,M.A., Zatz,M., Zeldanis,J. and Setubal,J.C.

TITLE Submitted (02-JUN-2000) Organization for Nucleotide Sequencing and Analysis, Bioinformatics Lab - IC/Unicamp, C.P. 6176, Campinas, SP 13083-970, Brazil

JOURNAL

FEATURES  
source location/Qualifiers  
1. 51158  
/organism="Xylella fastidiosa"  
/db\_xref="taxon:2371"  
/plasmid="pXF51"  
41. 301  
/gene="XFa0001"  
41. 301  
/gene="XFa0001"  
/note="identified by sequence similarity; putative; ORF located manually"  
/codon\_start=1  
/transl\_table=1  
/product="transcriptional regulator"  
/protein\_id="AAF8570.1"  
/db\_xref="GI:9112239"  
/translation="MPMPHKQTVNKKVKAVERELIFGRFNNANMAKLSQREIAEK TGLTQNMISMVNGRSTSIDMAKLACVNPWLKLPLP"  
435. 1010  
/gene="XFa0002"  
435. 1010  
/gene="XFa0002"  
/note="identified by sequence similarity; putative; ORF located manually"  
/codon\_start=1  
/transl\_table=1  
/product="conjugal transfer protein"  
/protein\_id="AAF8571.1"  
/db\_xref="GI:9112240"  
/translation="MDFIALSQECPAIPWATOTMAITVTEGSPNLLISVNGAKLIVR OPEKRAEAVSTARMLIEKGYNDLIGIQVNSNLYKVGISVEDAFEPCKNLAAATIL QGNFQVAKGVHDQALHAALSAVNTGSFSGFNGVYQVKLNNAGVTVQVSQAQA VTKAKGEKASPOGSGVNVDTNTDVVY"  
1012. 2994  
/gene="XFa0003"  
1012. 2994  
/gene="XFa0003"  
/note="identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder"  
/codon\_start=1

gene  
CDS  
/transl\_table=1  
/product="topoisomerase I"  
/protein\_id="AAF8572.1"  
/db\_xref="GI:9112241"  
/translation="WKMTLIESPKIKKSTLIIDGCKIAASGVHADLPQKENGVA PDFPSYELTERGGVAVAKKNLKKQSDSYLADPDRBEGALSMHLQOQCKSNPLR VTFEIRIYANVKSILVAPRTIDVRVAAOERARLADLVGMSPELSRLTGKLSG ROSAPAVLAVVEREIKAFKVTNHPGAMFADAKTSGEWSAEMTLRPFVADNPY EMDDFAAVAVKAVVVKSPDESEAKRSPPTTSMOQAAVTLGVDPKAMDPA OKLYEOGHITVYRTDNNVSDGSDIYAVSVKGLILAEKPRKFA DAGOVGHPAV TPTWYEBEAGETSQKALYKLTILRALIACLDARAVRTVLEAOPQKNGVEE FRTGTLIYQGLKLIAGDQTEEDGKESNSNLPVCAPEGCLDVAHGKCKRTPR SHYQALVYKGLIEBEGIRPATVYAIWINDVSRGVYTKKYLEPTTGLIIVDSLG KFEELVGTREIESLDIRLAQSGSYKAVIYAKHDQIMNELSTLOVSAATPKPCOC GKPRRIIPGKNGHSGSGYDSCVTPDEGKGGOKVPEVSNFACVCGNPLKHLV KKGGGYDFMGCSGFKGECATYEDKKGNFEAK"  
3065. 3328  
/gene="XFa0004"  
3065. 3328  
/gene="XFa0004"  
/note="hypothetical protein; identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder"  
/codon\_start=1  
/transl\_table=1  
/product="hypothetical protein"  
/protein\_id="AAF8573.1"  
/db\_xref="GI:9112242"  
/translation="MNTISEELDPFKNAVKRLHIIONDAKEYEIPATLTWKEGDW HVTTRGKREWASLDRLSHIRKYSTLPISLTIRNLISK"  
3437. 3709  
/gene="XFa0005"  
3437. 3709  
/gene="XFa0005"  
/note="identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder"  
/codon\_start=1  
/transl\_table=1  
/product="conjugal transfer protein"  
/protein\_id="AAF8574.1"  
/db\_xref="GI:9112243"  
/translation="MKVSALCLTRAVFTPEIRPAAAAANDVAAQVLAITLGLRTTV IIGVACIGIAIAGLSWDMAIKITIGIVLIFGSASIVDYISAVA"  
3712. 4080  
/gene="XFa0006"  
3712. 4080  
/gene="XFa0006"  
/note="identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder"  
/codon\_start=1  
/transl\_table=1  
/product="conjugal transfer protein"  
/protein\_id="AAF8575.1"  
/db\_xref="GI:9112244"  
/translation="WNSNTQDVYTDPLFVGLITRPASIGIYEACVAFPMAGLIFLS VNNPLYLALPMHVAIVLISANDPKIFSSIHMLKTNRCRNLFWRASSFSPLATK KLEKRLTLRIYKXFRVEN"  
4094. 6541  
/gene="XFa0007"  
4094. 6541  
/gene="XFa0007"  
/note="identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder"  
/codon\_start=1  
/transl\_table=1  
/product="conjugal transfer protein"  
/protein\_id="AAF8576.1"  
/db\_xref="GI:9112245"  
/translation="MGAFAKYADALKNEKSTALFIPYSSHYTKVNVKLVNGDYATIR MGAHRESADYODINSWMDOLNGMRNIRASNTVWSHVVRREGEPPGGEYQPGFCH DNEKIRKHAHSKRLVNEIYUTVYTRQPKAVAKFELIPSSKAKELQOQOTBELRV VEDLISVALSIDRYEPLICGYEKGMSVEVLEFLAFVLDGEMRRKPLPRTETREIV LATSPPFGKGLMSLKGPTSTQFCAAVITIEYPSITPGILNELGMPFPEVLTQSF SFSKPAVAGRMKROQARMVNSGDVAGSIOIDEIDALDLTLTSNRFVGMVHFSMIRA

HDQKALNEHISDTGAVMSDSGMKMARADAGIAGAFWSQLPGNFEXRVAVGDITSRNFA  
 GFSSEHNFPISGRIRNNOWGDVAVMEKTTSGAPFYENFHKSEGDADAKSALDPMHKD  
 LANTVIGKSGIKRIVLEMLLAQTKESQAPFGKLSCEVLEKOLGAIAGVAMGAG  
 YYPKNGVPSGSENPROMPTTSNTLFTLILKOLVHESLPLTPROEIESEBIAAGVA  
 GAOKRIRIGALFELEDSTDSNGSLPRLERRCGGTTAMLDNPDITSLNCTTVGF  
 DVTEPLDNDARTPTITMILPHRILESILGRRIPIITDFEKLNDPATEDIAONCLVT  
 IRKODGFLVMTFOSPKOVKSPITAVIIGQDPATKILFNPEADDDVDVNGFKLEDER  
 EIVKSLGERSKQFLIKOGONSVALNLRGDEDELAIVSGNTATSLIAEKLVAEIGDP  
 PAVMLPBFHRIRKGAIA"  
 6538..7218  
 /gene="XFa0008"  
 6538..7218  
 /gene="XFa0008"  
 /note="identified by sequence similarity; putative; ORF  
 located using Glimmer/RBSfinder"  
 /codon\_start=1  
 /transl\_table=11  
 /product="conjugal transfer protein"  
 /protein\_id="AAE85577.1"  
 /db\_xref="GI:9112246"  
 /translation="MKLFKPPFGPALMIGLGVGNLANAGMPVIDMANTLEQAOLOVEA  
 MKQBOEMTSQIEBOYKQOYESTNGIRNNGSLVNNPAAQOVLPAVDATLLSQGVNMAA  
 IRSAEREDVMTSLAANSPTROAFQCAKQALANRAEMAYSTASRFSIDIVLLD  
 KINNPADAKDIAIDQARIQVEQVMQNEANFLQALQLLVNAQRDLQIQONKEISMKS  
 KGGMPFGW"  
 7339..7500  
 /gene="XFa0009"  
 7339..7500  
 /gene="XFa0009"  
 /gene="XFa0009"

## Alignment Scores:

Pred. No.: 546 Length: 51158  
 Score: 90.50 Matches: 47  
 Percent Similarity: 38.50% Conservative: 35  
 Best Local Similarity: 22.07% Mismatches: 80  
 Query Match: 8.09% Indels: 51  
 DB: 1 Gaps: 9

US-10-034-500-2 (1-218) x AEO03851 (1-51158)

QY 18 VALAAPTYYRserGlySerMetMetMetLeuValAlaValArgGluProLysIleGlu 37  
 DB 12575 GTTGAAGCTTCGGGCACTGTGACCCGTAATGAGAGTGAAGTCTG-----CTTGCT 12622  
 QY 38 LeuAlaLysGluAlaAlaIleLeuValIleAlaAlaMetProLysMetSerTyrglnGly 57  
 DB 12623 CTCAGAGAAAGCAATCCGATCAA-----GAATTTTGAAGCTCCCATGCAAC 12673  
 QY 58 GlyLeuTyrtThrPheAlaProTyrservalIleIleProGlnGlySerTrpAsnSerCys 77  
 DB 12674 GGCTTG-----ACGCTGGTGAATTTGTGAAAAACGGCTCGGGCAAAACCACT 12721  
 QY 78 ValAlaGluCysAlaValAsnThrIleLysSerAspLeuGluIlePheGlyArgLeuThr 97  
 DB 12722 ATCGGAAAAAGCATCAAAATTCATCCCAACAGAT-----GACGCGCTGTGTG 12769  
 QY 98 ProValGlyAspGlyIleLysMetIleGluThrValIleAsnGlnMetProProGlnAla 117  
 DB 12770 ACCGTTGAAGAC-----GTGATGAATAATGTTCTTGAACATCAACCCCAAGAGT 12820  
 QY 118 AlaValIleLeuLeuThrAspGlyIleAsnAsnLeuGlyMetAsnProValGluGluVal 137  
 DB 12821 CATCTGTTCTATTCGCGTGAAGGAGCAAGCTTAAGCATCAACCGAAGCAAGCATTA 12880  
 QY 138 LysSerIleTyrglnThrAsnProAsnValCysPheIleValIleSerPheAlaAspAsp 157  
 DB 12881 GCATGTGTCTAGAGATGAAGCTGACCGCATTTTGTGACGAAATGCGGGCGCATGAA 12940  
 QY 158 Ala-----Glu 159  
 DB 12941 GCATGGGAATTCGCAAGGCTGTGGTTCAGGCGATCCAGGATCAAGCAGGACACGCT 13000  
 QY 160 GlyLysAlaIle-----IleAspGlnIleValAlaLeuAsnSerGlySerValIleVal 177

DB 13001 GCGCGGGCGGTAGAAAGCCTTTGACCAGATGTGGCGCTTATCAAGATAGT-----GCT 13054  
 QY 178 AspGlyLeuGlnLeuLeuGlnAsnProAlaValCysGlnGlnPheValAsnSerValPhe 197  
 DB 13055 ACCGGCGGCAATTTG-----GACGCGCATAT 13081  
 QY 198 CysGlnGlnGlnIleLeuValThrGluGluValVal 210  
 DB 13082 ATCAAGAGCGAGTTTGTGACGCTGATATCGTCTT 13120  
 RESULT 32  
 ANGDHA  
 LOCUS  
 DEFINITION  
 Aspergillus nidulans gdhA gene for NADP-linked glutamate  
 dehydrogenase.  
 X16121  
 X16121.1 GI:2347  
 accession  
 X16121.1 GI:2347  
 version  
 X16121.1 GI:2347  
 keywords  
 Emericella nidulans.  
 source  
 Emericella nidulans  
 organism  
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 Eurotiales; Trichocomaceae; Emericella.  
 reference  
 1 (bases 1 to 2334)  
 authors  
 Hawkins,A.R., Gurr,S.J., Montague,P. and Kinghorn,J.R.  
 title  
 Nucleotide sequence and regulation of expression of the Aspergillus  
 nidulans gdhA gene encoding NADP dependent glutamate dehydrogenase  
 JOURNAL  
 Mol. Gen. Genet. 218 (1), 105-111 (1989)  
 MEDLINE  
 89384423  
 PUBMED  
 2550758  
 comment  
 Data kindly reviewed (25-FEB-1991) by Kinghorn J.R.  
 features  
 1..2334  
 location/Qualifiers  
 /organism="Emericella nidulans"  
 /db\_xref="taxon:162425"  
 /prim\_transcript 293..>1885  
 exon  
 293..445  
 /number=1  
 400..1885  
 /gene="gdhA"  
 join(400..445,499..774,828..1885)  
 CDS  
 /gene="gdhA"  
 /product="NADP linked glutamate dehydrogenase"  
 /codon\_start=1  
 /protein\_id="CA34252.1"  
 /db\_xref="GI:2348"  
 /db\_xref="SWISS-PROT:P18819"  
 /translation="MSNLVPEPEEQAYKELASTLENSTLFEQHPERYRALQVSVPE  
 RVIQPRVWENDGKGVQINRGYVQFNSALGPYGGIRFPSPVNLSTLKLAGFEOLF  
 NALTGLMGGGKGSDEDPKGSDEIRRCTAPMTLGHIGADDTLPAGDITGVTR  
 EVGLFQVRYRIRINQWEGVLTGKGSNGSLIRPANTGVVYVYVGMITHYVNGES  
 FAKRVAISGSQVAVQYALAKVLELGSVSLSDSKSLVKQESASFTEETIALID  
 LKVARQLSELSATSSAPAKFTYIPARPTNIPGKEVALPSATQNEVSGEAEHLI  
 KSGVRYIAEGSNMCTQALDIFEAHNNAPGDAIYVAPKAAAGVAVSGEAEHLI  
 SALSTLSEVEDRLKIMEDCFKNGLETKFATPAKGVLPVLVGSNIAGFTVAE  
 AKRQCGPMW"  
 mat\_peptide  
 join(400..445,499..774,828..1882)  
 /gene="gdhA"  
 /product="NADP linked glutamate dehydrogenase"  
 /EC\_number="1.4.1.4"  
 /note="glutamate dehydrogenase (NADP(+))"  
 446..498  
 /gene="gdhA"  
 /number=1  
 499..774  
 /gene="gdhA"  
 /number=2  
 775..827  
 /gene="gdhA"  
 /number=2  
 828..>1885  
 /gene="gdhA"  
 /number=3  
 exon  
 intron



```
/translation="MSQRSTKMGDTTAAIATASGAAGCITIRLGSGLITATGLMT  
TLRPRVAYHTRFLOVDEVIDGLALMPAPHSFGEDVELQHGSPLLRQLLRRL  
LDLQARQAHGPESEFERALNGKLDLIQAEALADIMGADLAAARAARSLSGVSRR  
EALAQOILRLRIHYEATIDPAESLDTRAOIRTSLOTLNSELVOLLRLDEHGRLC  
BALVTVLVGPPNVKSSLLNALIGSDRALVTDVGTDRITRESVHPHGLFEVVDPA  
GLREBGDAIEREGKRTINELQADALAVDADCPQIGSLADLALTSVRYMITIR  
KLDTLTPESALDVTIVPSAMTGAGETLRLKRLTLGTVETIRBEFARLRHVQ  
ALQRTAAHVTDANQAFAYEHLETAEBELRLAYKALGEINGSMSDBELGRJESNCIG  
K"  
complement (1868. .4579)  
/gene="XF2779"  
/complement (1868. .4579)  
/note="hypothetical protein; identified by sequence  
similarity; putative; ORF located using Glimmer/RBSfinder"  
/codon_start=1  
/transl_table=11  
/product="hypothetical protein"  
/protein_id="AAF85564.1"  
/db_xref="GI:9108025"  
/translation="MFKVTPLSFLLNLCTLALVACKHOSTPVESAPVTIVSDV  
KHSDDPOAOLPLASLHQLESYRIRIVLITDDEVOSFOERTSSOVGQILFNGLNO  
RTASRKIDALLASGAPSRPPTLVIIIDYESSPDYNDPLARLALDADAIND  
SALPAIKRQRIHEDIDALNOIERYNDETRLSIPPRNRGTFITKEKMDVYALIK  
RSYRKMLHDYGLLEAVPIEMSEHEIFENELPPTKLVLPDGPBRVYTNBKEI  
LOHYAVAPPEVGNRIGFRDRAKPOLGPIKRTLEBGYVANSVNMIDLSLK  
SGNALKEVINADIIILRAVDERRAPLPYGAASAEGLRISEIGLOSVMNDISD  
MADVPNSVARVLLOVARORGIILPHIDHRAVRLPOLLEKLAEGVPAAGMDG  
AFVNMTHPKAKTDGEGTERTASRNATVGVNDYTKPPLKTAANDAKOANTLLO  
SGPSSSHVILIKRRETRDKILSVFNDLAGRIQKNDRLVFRGHGATQSLRLR  
VGVIIPVDSDAHFAADDAISMTKVQNKAKTEAGHVALIMQACVSAIGLMKEITSP  
RLATISINMSROMKLSGGDDPVDKPNHGSVFTWALQASRAFDNDGLITGTE  
LASVYASAVTSVSAQTAPFGLGSGGSEFELPRRGEILITONTKRPVALINGP  
LESTNTROAKLHONTAAVKLITVKNLHDBEALLPAPVAKLKGKGAOPANLGLQ  
LYKXSKTELDIFSDTLELPHRLRTSKNGFPYPSQSEVAGHLPMLONTLKDHS  
ATTOLNKGDIYLOLNKDKARKTITVDVLOPEDSGTQGLKLTKL"  
complement (4670. .6367)  
/gene="XF2780"  
complement (4670. .6367)  
/note="similar to SP|P25754 (percent identity: 40 %/query  
alignment coverage: 89.2 %/subject alignment coverage:  
90.0 %); identified by sequence similarity; putative; ORF  
located using Glimmer/RBSfinder"  
/codon_start=1  
/transl_table=11  
/product="60kDa inner-membrane protein"  
/protein_id="AAF85565.1"  
/db_xref="GI:9108026"  
/translation="MNQTRVLLFSMLTVATLLMMDMGNKNETLEISASONLGDVSN  
LELHAYPOINAGAVPVQKDSQLIAVAPKVAVIVKTDVQLKIDGSPVLAADLRP  
OSKRGAKEPILKLTDDPRYPSATGWSQNSNPVNSTLPQPGVSYKLANDOR  
LVYPPVTTANGVSIKRTFTFERGRVILILDEIRNGSETWNAIVFRKSLVPTNI  
LNRMATNDSPSPNGAVVYSEKGYERAFADYNDGELNEIGGMVALLQNHFFA  
WIPKODASLYLQNGSRDIAELRGPATVAPQSTTTERLWVGLVQITKEHY  
KGLRVDVYSRFOALMIGGLFWILSHLSLNLHMGVALIGLVLLRIAYPLSAAQ  
KXSAKARKFOPLQOLKEKRGEDKROKQOALMKEKKEKINPMGGCPLIOMPIE  
ALVYVIVESVLEQLQAPMLGWTGIDLTDDPRVLTPLINLVIWMAVOKLTPPRADPTA  
GRMQVNPRLICVMAAFVPSGLALVYVINGLNLILQMMIRQHADPSRKSKSGNIK"  
complement (6412. .6840)  
/gene="XF2781"  
complement (6412. .6840)  
/note="similar to SP|P06277 (percent identity: 37 %/query  
alignment coverage: 74.6 %/subject alignment coverage:  
89.1 %); identified by sequence similarity; putative; ORF  
located using Glimmer/RBSfinder"  
/codon_start=1  
/transl_table=11  
/product="ribonuclease P"  
/protein_id="AAF85566.1"  
/db_xref="GI:9108027"  
/translation="MCLHLGTLNSCKRPSPRSARICLRSEYVVAEEQGRYSVLLRLH
```

```
HLTPSGVRVLGVSRVDIRAVNRNRIKALAEWKOIYKLVPGDYVVVROTAKD  
VSNALSELVSLSLRRIGALPLAIPIDNAMPFERNCRSK"  
complement (6878. .7018)  
/gene="XF2782"  
complement (6878. .7018)  
/note="similar to SP|P4370 (percent identity: 76 %/query  
alignment coverage: 93.5 %/subject alignment coverage:  
97.7 %); identified by sequence similarity; putative; ORF  
located using Glimmer/RBSfinder"  
/codon_start=1  
/transl_table=11  
/product="50S ribosomal protein L34"  
/protein_id="AAF85567.1"  
/db_xref="GI:9108028"  
/translation="MTRKTYQPSNLKRRKHGFRAMSTADGRKILARRAKGRKL  
SA"  
BASE COUNT 1980 a 1583 c 1643 g 1892 t 1 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 55.1 Length: 7099  
Score: 90.00 Matches: 54  
Percent Similarity: 38.37% Conserved: 45  
Best Local Similarity: 20.93% Mismatches: 101  
Query Match: 8.04% Indels: 58  
DB: 1 Gaps: 10  
US-10-034-500-2 (1-218) x AE004083 (1-7099)  
Qy 1 AAGUVALTRHAASeCyThrIyaaRyValGuseTyraenTyLeuValaAspTy 20  
Db 4246 GCACGCTTGCTCGTGGTGCCTGACGCTTTGATACCTTGACAGTAATGCTGATTAT 4187  
Qy 21 SerGlySerMetMet-----MetLyshVAlaValaArgGluProLysIle 36  
Db 4186 ATGAGTCTCCCTCGATCTGTATGACGCTGATCGTGGCTTTCGGAAGCGTTAAT 4127  
Qy 37 GluLeuAlaLysGluAlaIleLeuLysIleAsnAlaIleMetProLysMetSerTyGln 56  
Db 4126 GATCTT-----GACGCCGATTTAACCATATTCGGGCTTCCGGCATTAAGTTGGCT 4073  
Qy 57 GlyGlyLeuTy-----MetLyshVAlaValaArgGluProLysIle 60  
Db 4072 CAAGTATTCAAGAATATGATGATGATTAATGATGAGCGCACTACATGACGAA 4013  
Qy 61 -----ThrPheAlaProTy-----SerValIleIleProGlnGlySer 73  
Db 4012 ATTACTCGGCTATCTATTCCTTTTGATGCTGTAATCGGAACTTTATTAACGTAAGA 3953  
Qy 74 TrpAsnSerCyValAlaGluCyAlaValAsn---ThrIleLysSerAspLeuGluIle 92  
Db 3952 TGGAGTACGTAAGTGGCCGCAATTCGTCGCTTCAACACAGCTTAATGATTTTGATGAT 3893  
Qy 93 PheGlyLysLeuThr-----ProValGlyLysArgGlyIleLysMetGlnIleValIle 110  
Db 3892 TATGAAATTCGAAAGCTTATCCGATACCATGAGAAAGAAAGCAATGATATTGGT 3833  
Qy 111 AsnGlnMetProProGlnAlaAlaValIleLeuLeuThrArgGlyIleAsnAsnLeuGly 130  
Db 3832 AATGAATTAACCTCCTAACAACTTAGTACTTACCTTCGATGATGGCCCACTGAGTTTAT 3773  
Qy 131 MetAsnProValGluGluValLysSerLeTyGlnThrAsnProAsnValCyAspPheIle 150  
Db 3772 ACAATGAATCAAGAAATCTTACAGACATGCTGTG---CCAGCATGTTCTTGAAA 3716  
Qy 151 ValValSerPheAlaAsp-----AspAlaGluGlyLys----- 161  
Db 3715 GTTGAAGCTAATATTGGAGATGATGAGGCCCGAAACCGCAATTGACACTTATCT 3656  
Qy 162 -----AlaIleIleAspGlnIleValAlaLeuAsnSerGlySerValLeuVal 177  
Db 3655 AAATTAACCGTGAAGTAATAGACAAAGATACGACATGTTGTAATCATAGATGACCAT 3596
```

cy	178	AspGlyLeuGlnIleuLeuGlnAsnProAlaValCysGlnGlnupheValAsnSerValPhe	197
Db	3595	GACCTGTTACTTACCTAGTCGCACATCGCTCCGCCAAGAAGGTGATTAAATGCA-----	3542
cy	198	CysGlnGlnGlnIleLeuValThrGlnGlnValValValleuArgGlyValAsn	215
Db	3541	-----GATTAATTAATTTCCGCCGACTTGAT	3518
RESULT 34			
AE004357/c		17041 bp	DNA
LOCUS			linear
DEFINITION			BCT 31-JUN-2000
ACCESSION			Vibrio cholerae chromosome II, section 14 of 93 of the complete
VERSION			chromosome.
KEYWORDS			AE004357 AE003853
ORGANISM			AE004357.1 GI:9657547
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCES			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCES			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCES			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCES			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCES			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCES			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCES			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCES			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCES			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCES			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCES			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCES			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCES			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCES			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCES			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCES			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCES			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCES			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCES			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCES			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCES			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCES			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCES			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCES			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCES			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCES			
AUTHORS			
TITLE			
JOURNAL			

gene  
CDS  
/translation="MKKKKNIITKPTQGLPGHSAIDLEQALDOLANNRGOLAGLIAAPFPA  
PADIQROIEPPEQSTLANTMETETWACNSQHHIKDEYPPFLAMWESHQSPVAILKSLP  
NNHLDLILKEKEPMTWQQQGFKALDMLTLCPCPWLHLRPLAKPKANRYVAALDPI  
TTEQOAKALRLLELSRSIANTCDSLHIVSCWEYLENLYLHDQVILQVEQOILAE  
IAQKIRHEQALROLIDESGITGELVILHHLGKDELLIPMWVQENIDILLMGTLVAT  
GIAGPVGNTKMENTVQSIKESLVLPKKEGFVSPKNE"  
complement(1513) . 2748)  
/gene="VCA0160"  
complement(1513) . 2748)  
/note="similar to GB:J142023 SP:P44614 PID:1003469  
PID:1222213 PID:1204542; identified by sequence  
similarity: putative"  
/codon\_start=1  
/transl\_table=1  
/product="cryptophan-specific transport protein"  
/protein\_id="AAF96073.1"  
/db\_xref="GI:9657550"  
/translation="MKKPDIMSQKSLFGGACIIASVCVAGAGMLGPSAGAGWTLTLM  
SMLALMALWMTLSGMMLEAFKHYLDLRSPNTYCDMGLSHIRNRNNLITVYVGI  
LLVYITSSGILIDDLHINSKIASILFVAVFAFWHSTRVADRISVILIVENVSF  
LIFVSGALINVKTILPDTLNQSEYAPVAMAPVALSTSGHSAVSMSRAYGERR  
KAKVALIGETVIALSLVLMFSLIFGMLPRADFPAPVQQGNDVLLKALGSVSEK  
VSOAINASMAAILSSFGVGLGVDFADLPQSNCKQGRKTMVLVTPILSLILL  
PFGGITIAIGTGAATITKACTIPVLARKSRTLANAGQVVGGLALGVLVIFGV  
LTAVFHMMAMNLLPPEFG"  
complement(2827) . 4245)  
/gene="VCA0161"  
complement(2827) . 4245)  
/note="similar to GB:U10328 SP:P00913 GB:X15974 PID:1479997  
PID:290565; identified by sequence similarity: putative"  
/codon\_start=1  
/transl\_table=1  
/product="cryptophanase"  
/protein\_id="AAF96074.1"  
/db\_xref="GI:9657551"  
/translation="WENFKHLPEPPRIIRVIEPVKRTREYREKAILNAGNPELLDSE  
DVFILDLDSGTGAILTQEMQAMFRGDEAVSGSSYHALAAVADITGEYETITFHQG  
RGAEQIYLPVILKREKEKGLDRSKMVALSNYPDTTQGHQINCCVAKNYTTEADP  
TGKVDLPGNFDLEKLEQALIEAPNAVPIVSTITCNSAGQPSIANLKAVEILQ  
RYDIPVINDSRPAPENAYEITQORRDVNNMIEITBEAYAAAGNAAKMDAVOMQ  
GGLCFKNSSEFPDYTBERTICVQGEPTPYGLGGAEMYLAVGLDMQMDLAPV  
INQVELVNGSLAIGVITQAGHAAFDVAGKLLPHLPADQPPHAAACELYKAGIR  
AVEISLLGLGDPATGKQHPCPABILLTTPRATYQTHMDFIIIEAEKVANANWV  
GLEFPYBPVVRHETARLKERA"  
complement(4811) . 5017)  
/gene="VCA0162"  
complement(4811) . 5017)  
/note="identified by Glimmer2; putative"  
/codon\_start=1  
/transl\_table=1  
/product="hypothetical protein"  
/protein\_id="AAF96075.1"  
/db\_xref="GI:9657552"  
/translation="MNIVINDAGRLKRSVRKNIFLHDAEKNSGINVIHHPFLYFLI  
AVSORIHSDDDEWLSIIVKGLK"  
5026) . 5238  
/gene="VCA0163"  
5026) . 5238  
/note="identified by Glimmer2; putative"  
/codon\_start=1  
/transl\_table=1  
/product="conserved hypothetical protein"  
/protein\_id="AAF96076.1"  
/db\_xref="GI:9657553"  
/translation="MFEKTHILKLAQMKMPGKYAGRVLIDLPEBYLLMFOKQGFPEG  
ELGVLELCTIAIKIGESLIVPLKQ"  
complement(5308) . 5892)  
/gene="VCA0164"  
complement(5308) . 5892)



```

/gene="VCA0164"
/note="identified by Glimmer2; putative"
/codon_start=1
/translation="conserved hypothetical protein"
/protein_id="AAFG6077.1"
/db_xref="GI:9657554"
/translation="MRSVKSAMLTTPQLOVERRYLLKGAVCFHLSGPMRLPSAVLL
LIIFPNLSAQAQDLVKVSKRMRLQGNELIREYIALSGPKHGKQGDORTPE
GHVLLDFIEDSPFYSRHSISYPNTRDLQALQGVNPGDGIKIHGLKNGDEREGFV
OSPDWNCIAITNQMDEPLSLVQPGPILILEM"
6199..8331
/gene="VCA0165"
6199..8331
/gene="VCA0165"
/note="similar to PID:1653866; identified by sequence
similarity: putative"
/codon_start=1
/translation="GDEF family protein"
/protein_id="AAFG6078.1"
/db_xref="GI:9657555"
/translation="MFPRTLKRVTSNNRYMLAETPEYLTTELISQAGIVAPSO
DVAGHTLSSNTFTPLEAYPEAFMAAQAQPTSDGLIPAINTCRMNLPVWGSFI
FMLNHPQHRVTLIIQAACVDKVLHSTQSGELDFQLIAAQCLRAIEASKEFKR
DLRAQVYSEIROBQFIDNMKLHVQVLELSPNAPLDELHRAQVLEAMRHLDPSRA
LLLDPMKRCSFGTGDENGTIDEQTHYDLHOLEQYLEALSNECTLIKEVNDVP
LYTGVGVGGMAMLLIRGNDPTIGWIAIDNYINROITKOKOMESPSLIAQIY
IRKQOVNRMHASMELSRKMTYSEKSAVTPAIRMKIDRAVPLUTBACSYIO
CTWGIDQGNIVDSYFRGSTEINDIVLAVYPREVFKSVPIYHDKVGVGWA
MTMLTDGCTPIAFIAADNLIRRSPLTSQLEVINRMPANLEVLRAQAQIAIVLNE
TLELEVRNRTRDLQKANEKLDLMAKLDLTLGRNRMLEHOLEQCEQTIKEVNVGV
ILLDIDHGLFNNGCHLEGLDIALMRINILSRHASEHLEFCRIGSEFLLVANSR
ABEHLIAENIRKSIIEACIEHCENPSELTLVSGVNASYKREIOPDLYAEADK
ALYAKSGRQVGVIVENIDCIOAEM"
8577..8789
/gene="VCA0166"
8577..8789
/gene="VCA0166"
/note="similar to GB:123115 SP:P37410 GB:X52093
PID:349561; identified by sequence similarity; putative"
/codon_start=1
/translation="cold shock transcriptional regulator CspA"
/protein_id="AAFG6079.1"
/db_xref="GI:9657556"
/translation="MSQMTSSVKMFETKGFRTSDNGQDVFAHKSIVSSEFKT
LAEGRVSFVYEQKKGPOAAQVTL"
complement(9078..9866)
/gene="VCA0167"
complement(9078..9866)
/note="identified by Glimmer2; putative"
/codon_start=1
/translation="conserved hypothetical protein"
/protein_id="AAFG6080.1"
/db_xref="GI:9657557"
/translation="MVENVTRCGTSLSSVKRSINGRTSMNTKATVTLGSLIIGL
SALGFLVQQAQVKEKREYVAVAGLSREVAADVLPVIFTVADNQLSLFTVDQ
QTOITLQFLVEKQDRAISLAPAVIDKKAQYCEGDEABERATATQTLTVYSKQVD
VRKIISEIGLQKQGVNODPNNRVEFSFTGLNAIKPMIEATKQARBAQKFAK
DSQSLTGKIKTASQGFSTDRDNNTPYIKVIRVVTVEYLLSD"
complement(9955..10497)
/gene="VCA0168"

```

## Alignment Scores:

```

Pred. No.: 161 Length: 17041
Score: 90.00 Matches: 36
Percent Similarity: 42.96% Conservatve: 22
Best Local Similarity: 26.67% Mismatch: 53
Query Match: 8.04% Indels: 24
DB: 1 Gaps: 6

```

## FEATURES

## source

EMBL SC9395, Z46727 and at the end of this sequence by cosmid 8142.  
Location/Qualifiers  
1..19236  
/organism="Saccharomyces cerevisiae"  
/strain="AB972"  
/db\_xref="taxon:4932"  
/chromosome="IV"  
/map="4R"  
/clone="cosmid 9346"

```

US-10-034-500-2 (1-218) x AE004357 (1-17041)
QY 58 GtyleuYrTThrPheAlaProTySerValIle---IleProGlnGlySerTrpAsnSer 76
Db 14112 GACATGATCCGTGTTGGCCGATCAAGCTATCTGCACCAAGCCCTTGACCTGATGACAA 14053
QY 77 CyValAlaGluCyAlaValAsnThrIleLeuSerAspLeuGluIlePheGlyArgLeu 96
Db 14052 ACCGTCGCTAACCAAGCTTAATCAACAGCTA-----CTTAACCTATGCGCACTCA 14002
QY 97 ThrProValGlyAspGlyIleLeuSer---HisGluThrValIleAsnGlnMetProPro 115
Db 14001 ACCGCGATGATGGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 13942
QY 116 GlnAlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGly---MetAsnProVal 134
Db 13941 CAACCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 13882
QY 135 GluGluValIleSerIleTyrGlnThrAsnProAsnVal----- 147
Db 13881 GAAGCCGCGCATATGCGCAACACATGACCAACACAGATTTATACCGTGGCGTGGTGA 13822
QY 148 -----CysPheHisValIleSerPheAlaAspAla 158
Db 13821 GGTGAATGCTGCTCAAAATTTCTCTGTCAGTCAAGTGAATACCGCAAGATCTG 13762
QY 159 GluGlyIleAlaIleIleAspGlnIleValAlaLeuAsnSerGly 173
Db 13761 GATGAAAAGACACTG-----CAACCATGCGGACCAACCAAGT 13723
RESULT 35
SC9346/c
LOCUS SC9346 19236 bp DNA linear PLN 11-AUG-1997
DEFINITION S.cerevisiae chromosome IV cosmid 9346.
ACCESSION Z46784.1 GI:755782
VERSION Z46784.1 GI:755782
KEYWORDS CSB2; cytochrome b activator; MSB16; nucleoporin; RNA helicase;
SIR2 homologue; SLY1.
SOURCE Saccharomyces cerevisiae.
ORGANISM Saccharomyces cerevisiae.
REFERENCE 1 (bases 1 to 19236)
AUTHORS Oliver,K. and Harris,D.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 19236)
AUTHORS Barrett,B., Rajandream,M.A. and Walsh,S.V.
JOURNAL Direct Submission
TITLE Submitted (23-MAR-1995) Saccharomyces cerevisiae chromosome IV
JOURNAL sequencing project, Sanger Centre, Hinxton Hall, Hinxton, Cambridge
COMMENT CB10 IRQ B-mail: barrett@sanger.ac.uk
Notes:
All CDS over 100 codons have been analysed. CDS that are completely
overlapped and those that are overlapped by more than 50%
of their length by a larger CDS have been omitted from this
analysis.
Details of the omitted CDS are available on request. The more
significant matches with motifs in the PROSITE database are
also included but some of these may be fortuitous. The length in
codons and the calculated codon adaptation index (CAI)
is given for each CDS.
Cosmid 9346 is overlapped at the start of this sequence by cosmid
9355.

```



```

gene
/gene="SLY1"
<1..1466
/gene="SLY1"
/note="YD9346.01, incomplete SLY1 gene, len: 487, CAI:
0.18, SLY1_YEAST_P22213"
/codon_start=3
/product="SLY1p"
/protein_id="CAAB8703.1"
/db_xref="GI:755783"
/db_xref="SWISS-PROT:P22213"
/translation="LTNDPKTEETITGLCANIADGLFNTVLTINSIPITRAKGGP
AEIAEKGTKLRDVFINTNSSTSTLQGNLSERGVIIIDRNIDFAMNSHVIYO
CMVDFIFKLSRTVTIIPLESKENGDTNTAKPLATKTDIEBNDFMNSHLPPEA
ABNVEALNTKYEAEAITRKGTGNTSIDLONSNDTVOIOEVKKLPBELTAKNTI
DTHNNIFALILSLESKSLDPEFEVODRGSTGRPLDILKOGKTNNLSDKLSFI
VLTITSTGLPRDFOQVBNYBNTKENDYDINAKTYKRLREFQSLNSLQKSLSDGS
DSAEKPSNLITSLGIYGLTEGKQGVGSLISGIRKLPBKRTPIITNVADIMPLNS
SQKNLETTDSYLYIDPKITRGSHTRKPKRQSNKSLVFVVGNGVLEYQNLQEVASHQ
LHPKVMYGSALTTPPAFLNEISRLGASNSNDA"
1..1004
/misc_feature
/note="overlap with cosmid 9395"
/complement(1676..3069)
/note="YD9346.02c, unknown, len: 463, CAI: 0.19; PS00017
ATP/GTP-binding site motif A (P-loop)"
/codon_start=1
/product="unknown"
/protein_id="CAAB8704.1"
/db_xref="GI:755784"
/db_xref="SPTREMBL:Q03940"
/translation="WVAISEVKEPNVSNSSGAVTRPAATHIKGLGIDSEGVAKRV
EGGFGIOEARACGIVDLIKAKSGRAIILAGSGTGKALALAIISOELGPVFP
CPLVSGELSYVEVKEETLMEFRRAIGIRIKETEVEYGEVTELTPDAENPLGGG
KTIHSIVYGLSAKSTKTRIDPTIYBSIQREKVSIGVITYIEANTGAVKRGSDAY
ATEFDLEEEVPLPRGEVHKKEIIVQDTLHDLVANAPQGGQDVISMWQQLPKP
KTEITERKROEVNKKVAKYIDQVANELPGVAFIDBVNMLDIEFTYLNKALNSIAP
VVVLASNGMTVTRGTEVDVISPBGVPLDILIVRTLPHYKDEIRTLIERAAVER
LOVESSALDILATMGTESTSLRYALQLLAPCGILACTSNRKEIVNVDAEAKLFLDAK
RSTKILTSANVL"
/misc_feature
/complement(2812..2835)
/note="PS00017 ATP/GTP-binding site motif A (P-loop)"
3411..4523
/note="YD9346.03, unknown, len: 370, CAI: 0.13, low
similarity to SIR2, SIR2_YEAST_P06700 (26.9% identity in
301 aa overlap)"
/codon_start=1
/product="unknown"
/protein_id="CAAB8705.1"
/db_xref="GI:755785"
/db_xref="SWISS-PROT:P53688"
/translation="MKOKFVLPITPSTAEKKPOTENRCNENLKPRILLPOLKESVRN
RRPRISYRELNSVFDLDAVYDSTHLSKQRRHMDRDAFISYALNSTRMYVVGAG
ISVAAIGIDPFRSEGIFFSTVNGSGKDLVYNRVGDSEMSLKFQNLVSLFRLSNG
QPTKHEHMEAFARQRLRLTYQNTDGLDTQLPHLSTVPLAKPIPTVOLAGSIKH
MECNKLNKIPEDPELFKCDKDFDSTELIPSCPOEEYETVRKMAGLASTGVGKLRP
RVLYNVEVPEGDFEIGELIANNDLKKRIDCLIVTYSIKIPGVKNICRQPAAYVHANRG
IYLYNTSMPPNVLDLKFVLDLVVGDDQHTSL"
/misc_feature
/complement(4643..5935)
/note="19346.04c, unknown, len: 430, CAI: 0.10, low
similarity to nucleoporins"
/codon_start=1
/product="unknown"
/protein_id="CAAB8706.1"
/db_xref="GI:755786"
/db_xref="SWISS-PROT:P49686"
/translation="MSAFGNPTSGAKPLMNTSGINPTNNASTNNNGSAGFGRPS
FGTANMTGATTSAGFQPOFGTNGTNTGNTISAGNTSNAKPSAFAPAGSAP
IVNVPSTTSARFAGPSFGSTGAMATSPNGKSGMSGAGFGOPAFANTTAIPSS
IYNSNSNARFAGASNPPLTTSPFGSLQONASQNASSTSAFGKTPGKANTTQSPG
TTQNTSTSGTGVSPFGTGTNNKNSPFSNLQSGAGASSPFGITTSAAANNVNGS
SAFGTNNOSPFSGSGGTFGSASNLKNTNGTNGFOSFNGKSGFSGITIQDQANKVSK
SNPSFGQWPNTPDNISLKSNGNATSGFGQOQONMATVNAVATGKIRFVQGLSEK

```

```

gene
/gene="MS116"
/complement(7026..9020)
/gene="MS116"
/complement(17026..9020)
/note="YD9346.05c, MS116 gene, len:664, CAI: 0.21
MS16_YEAST_P15424 ATP-dependent RNA helicase MS116;
PS00017 ATP/GTP-binding site motif A (P-loop); PS00039
DEAD-box subfamily ATP-dependent helicases signature"
/product="MSs116p"
/protein_id="CAAB8707.1"
/db_xref="GI:755787"
/db_xref="SWISS-PROT:P15424"
/translation="MLTSIIKGTPTVLAASRNLLAALSNCHITWAVSRRLYNDGND
ORNGRQNRNNNSRNSRNSRSPRSREDDDEVHFDKTFSLIHPKEDNSKEY
TDLSELEKDEKIKHAKITRMEPRGLTPQOKTKIKLISEEDHDLVIAKGTGKTF
AFLLPIRQHLINTPFGQYKAVIYVATPRLDAQIYBAVKKIHDMMYQKKVAVSL
VGGDFDPAKNNKRLNPVIATPGLDIVLEKISNFRFRVDYKYLDEADRLLEIG
FRDLEITISGLNKNKSNADNITLLEFATLDKVOKLANNINNKKECLELDTVDN
EPEAHERIDQSVIASEKANSIFAVERHIKKOIKERDSNYKAIIPATVKTSLCSI
LKNFEKKDLPILEPHGKITONKRTSLVRFKDESGILVCTDVGARGDPFNHVEVLQ
IGVPSELANYHRIQRTARSGKEGSSVLFCKDLPRVLEDAKNTVIAKQCYEPS
EIKSYVLEATNEEPEDISDVILISYBSCICEVRFSERRLIETASTGVILNP
QKTIYVSRFLDKLGLSRSPIGKAMFEIRDTYSSRDGNKKSIDYDDDEISFRKANYN
NRSQNRDYDEPFRSRNNRRSFSRSDKNYSSRNSNTY"
/misc_feature
/complement(8202..8228)
/gene="MS116"
/note="PS00039 DEAD-box subfamily ATP-dependent helicases
signature"
/complement(8544..8567)
/gene="MS116"
/note="PS00017 ATP/GTP-binding site motif A (P-loop)"
9673..11274
/gene="REF2"
/9673..11274
/gene="REF2"
/note="YD9346.06, REF2 gene, len: 533, CAI: 0.15;
RNA-binding protein involved in mRNA 3'-end formation;
PS00599 Amino transferases class-II pyridoxal-phosphate
attachment site; conflict with GENBANK entry SCU20261;
accession U20261 at about 10949 results in a frameshift
with respect to this sequence"
/codon_start=1
/product="REF2p"
/protein_id="CAAB8708.1"
/db_xref="GI:755788"
/db_xref="SWISS-PROT:P42073"
/translation="WSAPVQLVNIASHLQASTIQOIRLDVNDPDKCKSSIQIARI
DKYIDSLQALNQPTKDNLIHERKEKNTVEDIOLYSGLSKMYLDIYQLIKLREKQ
HSHPTPIANDVSLDFVNLQLPKPSPEERKNYIDNLIKNKSNRLSKNDGIVDAVINL
CVLDTSVENVRSYKLLDITDGFQKSGSTGTKNLKKGLASSAKIIDSKEKEKEL
DKSYKVTATKPKPLANNNDKNSPSTASGMKKLKSGLFNKNKAKTESLEPTSS
KKKLSFSKYLKDDMDTKLSTKTSIDVDPRYNPEASIVANIISSSTSGSSTTVAT
PASSSEPIKKTIKLSVQDSNVQSLIRNKPKKARISITKFLDDSDILVYGDLPNOG
LOVSTPTQKLIKPEKEGEPKEIILFLEMGSILKFLPDLMPKANTNSDYMIOISFKGG
PIHCETPTPLYRKRNHFNPNDLNKRPREPIEDLNQNTNSTPTIAKAFKNLSLLR
KDRGGLPKVHPIVYKXKYPYRPHV"
11155..11184
/misc_feature
/gene="REF2"
/note="PS00599 Amino transferases class-II pyridoxal
-phosphate attachment site"
/complement(11348..12073)
/note="YD9346.07c, unknown, similar to T0SG5.5 CE00504
hypothetical protein A (T. aquaticus) (40.8% identity in
218 aa overlap); PS00017 ATP/GTP-binding site motif A
(P-loop)"
/codon_start=1
/product="unknown"
/protein_id="CAAB8709.1"
/db_xref="GI:755789"
/db_xref="SWISS-PROT:Q03941"
/translation="WLVLGLTGIAAGKSTVSRRLRDKYKLPYVDADKIAHQVBERGQ

```

NAVDQIVLYFKDKIPNLLLEDGHLNREALGKWFVSHKEDLQALNGITHPAIRYMFKE  
IGYVYKNGKMCVADVPILFEGNIDSLICGVSVISICQELDELRLMRLNPLSEDAK  
NRLNSQMSSTERMARSDYILIONNSNIDLVLYEQIESVVKIKPISKLRTLYEYFPFGAV  
SASSIVMSRLMKLONKSSAV"  
complement (12029..12052)  
misc\_feature  
/note="PS00017 ATP/GTP-binding site motif A (P-loop)"  
gene  
12302..13471  
/gene="CBS2"  
12302..13471  
/gene="CBS2"  
/note="yD9346.08, CBS2 gene (CBP7), len: 389, CAI: 0.13,  
CBS2\_YEAST P14905, cytochrome b translational activator"  
/codon\_start=1  
/product="CBS2p"

Alignment Scores:

Pred. No.: 186 Length: 19236  
Score: 90.00 Matches: 52  
Percent Similarity: 33.98% Conservative: 35  
Best Local Similarity: 20.31% Mismatches: 69  
Query Match: 8.04% Indels: 100  
Gaps: 12

US-10-034-500-2 (1-218) x SC9346 (1-19236)

Qy 16 TyrluValAspTyrSerGlySerMetMetMetLysHisValAlaValArgGluProLys 35  
Db 18760 TACATGTTGACTT-----CACATGTTGTGAGATTCACACAA 18722  
Qy 36 lIegluLualValyGluAlaIleLeuValIleAenAlaIleMetProLysMetSerTyr 55  
Db 18721 TTTCAAAAGGGTAAAGATTGTCCAGAAATACCGAGACTGAAATTTCTACTTATT 18662  
Qy 56 GlnglyGlyLeuTyrThrPheAlaProTyrSerValIleIleProGlnGlySerTrpAsn 75  
Db 18661 CAA-----TTCAGTAAGATATATACCCAC----- 18638  
Qy 76 SerCyValAlaGluCysAlaValAlaAnThrIleLysSerAspLeuGluIlePheGlyArg 95  
Db 18637 -----TTGAAAATATTTTGGAAATATCTGAATCA 18608  
Qy 96 LeuThrProValGlyAsp-----GlyIleLysMet----- 105  
Db 18607 CTTCAAGTAGCTACAGATCGACGAATTCGATCCAAATTTGGCGCCATGGAACCTT 18548  
Qy 106 -----HisGluThrValIleAsnGlnMetProProGlnAlaIleValIleLeuLeuThr 123  
Db 18547 TTGAACCAATTCCTAATATCTTTTACAAATCTCTCTAGAC-----CTGGTATTTC 18497  
Qy 124 AspGlyHisAenAsnLeuGlyMetAsnProValGluGluValIleSerIleTyrGlnThr 143  
Db 18496 CCAGAGATTAACAATCTTGGCTATG-----AAGGAATGTTCCAGAT 18455  
Qy 144 AsnProAsnValCys-----PheHisValIleSerPhe----- 154  
Db 18454 TGTACTCAGTCTCGAATCCACAGCTCAGATCTAGAGCTGGAACCTTACGTTGTAGG 18395  
Qy 154 ----- 154  
Db 18394 AATGAGCTTTGCAATGAACTACGAAATCTAATTAAGTGACTAAACCTTGTGCGAG 18335  
Qy 155 AlaAspAspAlaGluGlyValAlaIleIleAspGlnIle----- 167  
Db 18334 ATTGATATAAAACGCGAGGATCTTTGGCCAAATTAAGAAATCAAGTGAATGA 18275  
Qy 168 -----ValAlaLeuAsnSerGlySerValLeuValAspGlyLeuGln 181  
Db 18274 AGAAACAAACTTATCTAAGATCTCTCAGAAAACGGGTACAGTCCAGGATCCACA 18215  
Qy 182 LeuLeuGlnAsnProAlaValCysGlnGluPheValAsnSerValPheCysGlnIleGln 201  
Db 18214 TTACTTAACCAAC-----ATAATTTCTTTTCAAAAGTGAAGCT 18176

Qy 202 lIeLeuValThrGlu--GluValValLeuAspGlyValAsnPhe 216  
Db 18175 ATAACTTCCAGAGCTCAGGAATATTATTAAGAGGCGCTACTTTC 18128

RESULT 36  
AC114131  
LOCUS  
DEFINITION  
AC114131 101054 bp DNA linear HTG 13-JUL-2002  
Rattus norvegicus clone CH230-131G8, \*\*\* SEQUENCING IN PROGRESS  
\*\*\*, 56 unordered pieces.  
AC114131  
AC114131.2 GI:21735299  
VERSION  
KEYWORDS  
HTG; HTGS PHASE1.  
SOURCE  
Norway rat.  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE  
AUTHORS

1 (bases 1 to 101054)  
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-ouman,F.R., Allen,C.,  
Albrooks,S.L., Amarantunga,H.C., Are,J.R., Ayale,M., Banks,T.,  
Barbaria,J., Benton,J., Bimarge,K., Blankenburg,K., Bonnin,D.,  
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,  
Bunay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
Davis,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
Einhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escoto,M.,  
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Franz,P.,  
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Gortell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,K., Hart,M., Havlik,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,  
Homsli,F., Howard,S., Huber,J., Huliy,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Loughesed,H.,  
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheshwari,M., Mapua,P., Martin,R., Mattindale,A., Martinez,E.,  
Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,  
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
Moser,M., Neal,D., Newton,J., Newson,N., Nguyen,A., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogdu,M., Okunnu,G.,  
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
Peterson,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
Rivers,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G.,  
Scherer,S., Scott,G., Shen,H., Shoohtari,N., Slason,I.,  
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,  
Stutson,A., Swatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansley,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
Umani,K., Vasquez,L., Vera,V., Villalob,D., Vinson,R., Wang,Q.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,  
Williams,G., Williamson,A., Wleczky,R., Wooden,S., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstock,G. and Gibbs,R.  
Direct Submission  
Unpublished  
2 (bases 1 to 101054)  
Worley,K.C.  
Submitted (07-MAR-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 101054)  
Worley,K.C.  
Direct Submission  
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
COMMENT  
On Jul 12, 2002 this sequence version replaced gi:19224458.

```

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GNL
Center clone name: CH230-131G8
----- Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 38071 bases at least Q40
Consensus quality: 40739 bases at least Q30
Consensus quality: 43197 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 56 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 1074 1073: contig of 1073 bp in length
* 1174 1173: gap of unknown length
* 2405 2404: contig of 1231 bp in length
* 2505 2504: gap of unknown length
* 3685 3684: contig of 1180 bp in length
* 3785 3784: gap of unknown length
* 4985 4985: contig of 1201 bp in length
* 5086 5086: gap of unknown length
* 6300 6299: contig of 1214 bp in length
* 6400 6399: gap of unknown length
* 7417 7416: contig of 1017 bp in length
* 7517 7516: gap of unknown length
* 8750 8749: contig of 1233 bp in length
* 8850 8849: gap of unknown length
* 10213 10212: contig of 1363 bp in length
* 10313 10312: gap of unknown length
* 11334 11333: contig of 1021 bp in length
* 11434 11433: gap of unknown length
* 12895 12895: contig of 1462 bp in length
* 12986 12985: gap of unknown length
* 14180 14179: contig of 1184 bp in length
* 14280 14279: gap of unknown length
* 15720 15719: contig of 1440 bp in length
* 15820 15819: gap of unknown length
* 16861 16860: contig of 1041 bp in length
* 16961 16960: gap of unknown length
* 18167 18166: contig of 1206 bp in length
* 18267 18266: gap of unknown length
* 19662 19661: contig of 1395 bp in length
* 19762 19761: gap of unknown length
* 21070 21069: contig of 1308 bp in length
* 21170 21169: gap of unknown length
* 22424 22423: contig of 1254 bp in length
* 22524 22523: gap of unknown length
* 23945 23945: contig of 1422 bp in length
* 24046 24045: gap of unknown length
* 25373 25372: contig of 1327 bp in length
* 25473 25472: gap of unknown length
* 26873 26873: contig of 1511 bp in length
* 26984 26984: gap of unknown length
* 27084 27083: contig of 1463 bp in length
* 28547 28546: gap of unknown length
* 28647 28647: contig of 1544 bp in length
* 30191 30190: gap of unknown length
* 30291 30291: contig of 1376 bp in length
* 31667 31666: gap of unknown length
* 3167 31667: contig of 1781 bp in length

```

```

* 33548 33647: gap of unknown length
* 33648 33648: contig of 1648 bp in length
* 35296 35295: gap of unknown length
* 35396 35395: contig of 1998 bp in length
* 37394 37393: gap of unknown length
* 37494 37493: contig of 1638 bp in length
* 39132 39131: gap of unknown length
* 39232 39231: contig of 1396 bp in length
* 40628 40627: gap of unknown length
* 40728 40727: contig of 1837 bp in length
* 42565 42564: gap of unknown length
* 42665 42664: contig of 1750 bp in length
* 44415 44414: gap of unknown length
* 44515 44514: contig of 1444 bp in length
* 45959 45958: gap of unknown length
* 46059 46058: contig of 1672 bp in length
* 47731 47730: gap of unknown length
* 47831 47831: contig of 2421 bp in length
* 50252 50251: gap of unknown length
* 50352 50351: contig of 1960 bp in length
* 52312 52311: gap of unknown length
* 52412 52411: contig of 1475 bp in length
* 53887 53886: gap of unknown length
* 53987 53986: contig of 1443 bp in length
* 55430 55429: gap of unknown length
* 55530 55529: contig of 2557 bp in length
* 58087 58086: gap of unknown length
* 58187 58186: contig of 1657 bp in length
* 59844 59843: contig of 1657 bp in length
* 59944 59943: gap of unknown length
* 62193 62192: contig of 2249 bp in length
* 62293 62292: gap of unknown length
* 64236 64235: contig of 1943 bp in length
* 64336 64335: gap of unknown length
* 66151 66150: contig of 1815 bp in length
* 66251 66250: gap of unknown length
* 67694 67693: contig of 1443 bp in length
* 67794 67793: gap of unknown length
* 69007 69006: contig of 1213 bp in length
* 69107 69106: gap of unknown length
* 71094 71093: contig of 1987 bp in length
* 71194 71193: gap of unknown length
* 73580 73579: contig of 2386 bp in length
* 73680 73679: gap of unknown length
* 74957 74956: contig of 1277 bp in length
* 75057 75056: gap of unknown length
* 75057 75056: contig of 2102 bp in length
* 77159 77158: gap of unknown length
* 77259 77258: gap of unknown length
* 79237 79236: contig of 1978 bp in length
* 79337 79336: gap of unknown length
* 81082 81081: contig of 1745 bp in length
* 81182 81181: gap of unknown length
* 83323 83322: contig of 2141 bp in length
* 83423 83422: gap of unknown length
* 84423 84422: gap of unknown length
* 84885 84884: contig of 1462 bp in length
* 84985 84984: gap of unknown length
* 86416 86415: contig of 1431 bp in length
* 86516 86515: gap of unknown length
* 90143 90142: contig of 3627 bp in length
* 90243 90242: gap of unknown length
* 90243 94508: contig of 4266 bp in length

```

```

Alignment Scores:
Pred. No.: 1,41e+03 Length: 101054
Score: 90.00 Matches: 50
Percent Similarity: 42.27% Conservative: 43
Best Local Similarity: 22.73% Mismatches: 81
Query Match: 8.04% Indels: 46
DB: Gaps: 9

```

```

US-10-034-500-2 (1-218) x AC114131 (1-101054)
OY 3 ValThAlSeCysThrIys-----ArgValGlusEtyrAsnTyrlEuValAaPtyr 20
:::|||||
|||||
:::

```

```

Db 53247 ATTACTTAAATATACCAAGTTGCATCAGTTGCTCTACCCACCTACTTACTTACATC 53306
Qy 21 SerGlySerMetMetMetLysHisValAlaValArgGluProLysIle----- 36
Db 53307 TCTAGCTGAATGCAGAAATGATATCTCAAACTCCCAAGTACCAGTAAATTCCTAATTC 53366
Qy 37 ---GluLeuAlaLysGluAlaIleLeuLysIleAsnAlaAlaMetProLysMet----- 53
Db 53367 CAAAGACTGCAGATCAATGATCTCTCTCCAGTGAACAGCAACATTCCTTAGAAG 53426
Qy 54 -----SerTyrGln-GlyGlyLeuTyrThrPheAlaProTyrSerValIleIleProG 71
Db 53427 GCACCAAAATTCAGTGCATGCTCTATTCATTCCTCAATATTCCTC----- 53472
Qy 71 nglySerTrpAsnSerCysValAlaGluCysAlaValAsnThrIleLysSerAspLeuG 91
Db 53473 -----CTCTCACTACTCTACAGAAATATCTGTCAGGTCCACCTTAGA 53513
Qy 91 uilePheGlyArg-----LeuThrProValGlyAspGlyIleLysMetHisGluThrVa 109
Db 53514 AGTAGCTGAAATCAGTAAATCTCTCCCTGCTCGTGAATATATCACAGATCCAAACCAT 53573
Qy 109 ILLAsnGlnMetProProGlnAlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLe 129
Db 53574 TGCCACTTGCAATG-----CTGAAATATACGAAATGTCACACACAGCAT 53615
Qy 129 u-----GlyMet-AsnProValGluGluValLysSerI 140
Db 53616 CAACCTGCTTTTCCATACACCCCTACATGAAATGTTCTCCGCTCACACAGTAAATAA 53675
Qy 140 leTyrGlnThrAsnProAsnVal-----CysPheHisValSerPheAlaAspA 157
Db 53676 CATGTCACCTCATGCAACTTGTTCATCAAGTCCCTTCATGATCTTACTTACAGAGCAG 53735
Qy 157 apAlaGluGlyValAlaIleIleAspGlnIleValAlaLeuAsnSerGlySerValLeuV 177
Db 53736 AAGCCAAAGGCCCCCAACCATCCTTTAGAGTCTC-----CTATAGACTTATACATCTTAT 53789
Qy 177 aLAspGlyLeuGlnLeuLeuGlnAsnProAlaValCysGlnGluPheValAsnSer 195
Db 53790 ACGACTCTCTCTCCGCTTACTCTCTCCCAAGCCACACTGCGCTCCTGCTATCTTCA 53845

RESULT 37
AL591983/c 324050 bp DNA linear BCT 06-JUN-2002
LOCUS Listeria monocytogenes strain EGD, complete genome segment 11/12.
DEFINITION AL591983 AL591824
ACCESSION AL591983.1 GI:16411809
VERSION
KEYWORDS
SOURCE
ORGANISM
Listeria monocytogenes.
Listeria monocytogenes
Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
REFERENCE
1 Glaaser, P., Frangeul, L., Buchrieser, C., Rusniok, C., Amend, A.,
  Baquero, F., Berche, P., Bloeker, H., Brandt, P., Chakraborty, T.,
  Charbit, A., Cherouani, F., Couve, B., de Darvar, A., Dehoux, P.,
  Domann, B., Dominguez-Bernal, G., Duchaud, E., Durant, L.,
  Dussauguet, O., Entlian, K. D., Feibi, H., Portillo, F. G., Garrido, P.,
  Gautier, L., Goebel, W., Gomez-Lopez, N., Hain, T., Hauf, J.,
  Jackson, D., Jones, L. M., Kaerle, U., Kretz, J., Kuhn, M., Kunst, F.,
  Kuraprat, G., Madueno, E., Maitournam, A., Vicente, J. M., Ng, E.,
  Medjari, H., Nordstiek, G., Novella, S., de Pablo, B., Perez-Diaz, J. C.,
  Pucell, R., Remmel, B., Rose, M., Schluener, T., Simoes, N.,
  Tlerriz, A., Vazquez-Bolaji, J. A., Voas, H., Wehland, J. and Cossart, P.
  Comparative genomics of Listeria species
  Science 294 (5543), 849-852 (2001)
JOURNAL
MEDLINE
21537279
PUBMED
11679669
REFERENCE
2 (bases 1 to 324050)
AUTHORS
Glaaser, P., Frangeul, L. and Rusniok, C.
TITLE
Direct Submission
JOURNAL
Submitted (06-JUN-2001) Glaaser P., Institut Pasteur, Genomique des
  Microorganismes Pathogenes, 25 rue du Docteur Roux, 75724 Paris

```

```

COMMENT
E-mail: pglaser@pasteur.fr
Phone: +33 1 45 68 89 96, Fax: +33 (0) 1 45 68 87 86.
FEATURES
source
1. 324050
   /organism="Listeria monocytogenes"
   /strain="EGD-e"
   /db_xref="taxon:1639"
   /complement(1290..492)
   /gene="lmo2321"
   /complement(1290..478)
   /gene="lmo2321"
   /codon_start=1
   /transl_table=1
   /protein_id="CAD00399.1"
   /db_xref="GI:16411810"
   /db_xref="SPTREMBL:Q928W7"
   /translation="MAERIFRKQTFPNSEIFIDRTKMTANPAFRQKIALITGCEK
MTDYIELKIKGYEYVR"
   /complement(487..492)
   /gene="lmo2321"
   /complement(586..801)
   /gene="lmo2322"
   /complement(586..801)
   /gene="lmo2322"
   /codon_start=1
   /transl_table=1
   /product="gp44 [Bacteriophage A118]"
   /protein_id="CAD00400.1"
   /db_xref="GI:16411811"
   /db_xref="SPTREMBL:Q8Y4V8"
   /translation="MSEVLVASVYEGYESKINFTENINNIWKEFKKXIDEVERKKRAE
TFNKQYVTEKELVDGHLREIIIPRT"
   /complement(1454..1331,1340..1345)
   /gene="lmo2323"
   /complement(798..1331)
   /gene="lmo2323"
   /codon_start=1
   /transl_table=1
   /product="gp43 [Bacteriophage A118]"
   /protein_id="CAD00401.1"
   /db_xref="GI:16411812"
   /db_xref="SPTREMBL:Q8Y4V7"
   /translation="MFQKSTSPAAQVLAETRTQKELAIDSYTPALISNQIGKRT
VSLQAEQLIDSVNEPSTYLPFAHEPENGMIPLPFDGLDHHHSLTRFELEVEAAN
TKRGLETPNLRKSGMLQREAAKQAISETTDVVASALTINSAIALAFNIDQOVLAN
KRDIYQKSGLVNSCGK"
   /complement(810..815)
   /gene="lmo2322"
   /complement(810..815)
   /gene="lmo2322"
   /complement(1340..1345)
   /gene="lmo2323"
   /complement(1454..2230)
   /gene="lmo2324"
   /complement(1454..2230)
   /gene="lmo2324"
   /note="similar to anti-repressor [Bacteriophage A118]"
   /codon_start=1
   /transl_table=1
   /protein_id="CAD00402.1"
   /db_xref="GI:16411813"
   /db_xref="SPTREMBL:Q8Y4V6"
   /translation="MSKIQYIANEMLPVLENEQEKRPVDAEMLEKLIANTRPNDWY
RMGNQYENGLEDFYSLTKNGRSPSTNYFLTDITAKELMANVONNENGRAIRKPIEV
EKARQKATPEPAFSWIEDPVARAKWIEQOEKQVLKLEQKRVFAEAVQTS
ENTLIVDLATILKQGLDIGQNRLPFEMLRGSGYLKSGAYVNNPDSQKAVNLGIFEOK
THIHTDNGMLKTTYPQVTKGQVYLLNKLLEBNQVIT"
   /complement(2212..2487)
   /gene="lmo2325"
   /complement(2212..2487)
   /gene="lmo2325"
   /codon_start=1

```



Db 120911 GTAGATCTTCCAGCAATGGCAGATGAAACAACACTATGCAAGTTTCAGTACGGAACCAACA 120852  
Qy 174 SerValLeuValAspGlyLeuGlnLeuLeuGln-----AsnPro 186  
Db 120851 AGTGAACATTAATGAGGGGAGATGACTTAAGACAGATACATTAGATCATTTAAACGAA 120792  
Qy 187 AlaValCysGlnGluPheValAsnSerVal-----PheCysGlnGluGlnLe 202  
Db 120791 GCAACACACTGTTACTATCATGTAACAACTTAACATGTAAGTAAGACAGACAGA 120732  
Qy 203 LeuValThrGluGluValValLeuArgGlyValAsnPheAla 217  
Db 120731 AAAGATGCCAAACAGCAATCTGTAATGCTGTAACCAATGCA 120687  
RESULT 38  
BC004727 2394 bp mRNA linear ROD 07-AUG-2002  
DEFINITION Mus musculus, similar to loss of heterozygosity, 11, chromosomal region 2, gene A, clone MGC:6690 IMAGE:3582797, mRNA, complete cds.  
ACCESSION BC004727  
VERSION BC004727.1 GI:13435731  
KEYWORDS MGC.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 2394)  
REFERENCE Straubeberg, R.  
TITLE Direct Submission  
AUTHORS Submitted (21-MAR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
COMMENT Contact: MGC help desk  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Gilbert Smith, Ph.D.  
Clava Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
Gunnarone, P.H., Garcia, A.M., Lu, X., Huylk, S.W., Hale, S.M., Yoon, V.S., Kowals, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAX Plate: 9 Row: d Column: 16  
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.  
Location/Qualifiers  
1..2394  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/map="FVB/N"  
/clone="MGC:6690 IMAGE:3582797"  
/tissue\_type="Mammary tumor. Metallothionien-TGF alpha model. 10 month old virgin mouse. Taken by biopsy."  
/clone\_lib="NCI CGAP\_Mam"  
/lab\_host="DH10B"  
/note="Vector: PCMV-SPORT6"  
347..1681  
/codon\_start=1  
/product="Similar to loss of heterozygosity, 11, chromosomal region 2, gene A"  
/protein\_id="AA04727.1"  
/db\_xref="GI:13435731"  
/translation="MEDAVKRVKALKKNIAGTEIITPLCNITKASIPGHPLOLFVPT

DGEVSDTFPSVIREVKNSKGRCPFSFGIGQASTSLIKNIARVSGTAIVITGDRMD  
TKALGSILKPLAQCAVDNIISLWDLPGLSVMSPEDLTFRCGLRIIVQTLMPQ  
VESTGAVCLHILQGRSLLENRVTSLOPKANDNPTIRHLAKSLIOTRDFSGOASKE  
EKEDVNNISLQSGVLSFTAFIATNKEDNIPVQSLAHARVPRVVMGSSMRVYSSP  
SGGKFGQPSRLRPAYDLCASBANLVTKKSASAIQKKTTNSTYKSNIKKHAKF  
GENAVVOLISLQKNGSWEIDEDILKILGTGSKOIKAANPKHEDPSAMATVLAIVWL  
HANQTDKCEWELLERRAVAMVLDHDAQSSIFMLVQANSLILKLSVNPVFGV"  
BASE COUNT 690 a 518 c 542 g 644 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 16.5 Length: 2394  
Score: 89.50 Matches: 55  
Percent Similarity: 43.65% Conservative: 31  
Best Local Similarity: 27.92% Mismatches: 88  
Query Match: 8.00% Indels: 23  
D: 10 Gaps: 11  
US-10-034-500-2 (1-218) x BC004727 (1-2394)  
Qy 2 GluValThrAlaSerCysThrIleValGluSerTyraNtyrLeuValAspTyrSer 21  
Db 113 GAAGTGAAGCCTCA-----AAGGCTGTGAGATTTGTCTTCATGACGCTTCA 166  
Qy 22 GlySerMet-----MetMetIleHisValAlaValArgGluProIleGluLeuAla 39  
Db 167 GGAATATGACCTCCCATGAGACAGAGAAACATTCGACGCTACGATGAGCGTCC 226  
Qy 40 LysGluAlaIleLeuIleLeuIleAsnAlaMetProIleMetSerTyr----- 55  
Db 227 AAGGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 286  
Qy 56 ---GlnGlyIleTyr---ThrPheAlaProIleSerValIleIleProGlnIleSer 73  
Db 287 GGAATGATCTTCTCTATGAGAAATTTCTTCGAGAGGTGAGATGACCTGAGATACA 346  
Qy 74 TrpAsnSerCysValAlaGluCysAlaValAsnThrIleLeuSerAspLeuGluIlePhe 93  
Db 347 ATGAGAGACGACGTAAAGA-----GTGAAGCTTTAAAGCAACTTAGGGGCTACT 400  
Qy 94 GlyArgLeuThrProValGlyAspGlyIleIleMetHisGluThrValIleAsnGlnMet 113  
Db 401 GAATCTTGACACCCCTGTCGACCAATTACAAAG-----GCATCTTCATTCCTGTCAT 454  
Qy 114 ProProGlnAlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsnPro 133  
Db 455 CCCCTACAG-----CTCTTGTCTTCACGAGATGAGAGATGAGACACATTAGTGTTC 508  
Qy 134 ValGluGluValIleYserIleTyrGlnThrAsnProAsnValCysPheHisValAlaSer 153  
Db 509 ATTGAGAGAGTTTAA-----TTAAACAGCAAGAAACACAGATGTTTC-----TCT 553  
Qy 154 PheAla---AspAspAlaGluGlyIleValIleIleAspGlnIleValAlaLeuAsnSer 172  
Db 554 TTTGGAATTGACAGAGAGCTTCGACGAGTTTAATCAAAATATGCCCCGGTATACAGG 613  
Qy 173 Gly---SerValLeuValAspGlyLeuGlnLeuLeuGlnAsnProAlaVal 188  
Db 614 GGTACTGAGTGTATTATCAAGGACAGACAGATGACAGACCAAGGCTCTT 664  
RESULT 39  
AL672070 197912 bp DNA linear ROD 29-JUN-2002  
LOCUS AL672070  
DEFINITION Mouse DNA sequence from clone Rp23-115a1 on chromosome 9, complete sequence.  
ACCESSION AL672070  
VERSION AL672070.12 GI:20339183  
KEYWORDS HTG.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 197912)

**AUTHORS** Bates, K.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (29-JUN-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk

**COMMENT**

On Apr 29, 2002 this sequence version replaced gi:20318531. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the submit sequences with corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) RP23-115A1 is from the RPI-23 Mouse PAC Library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>  
**VECTOR:** pBACe3.6

----- Genome Center  
 Center: UK Medical Research Council  
 Center code: UK-MRC  
 Web site: <http://mrcseq.har.mrc.ac.uk>  
 Contact: mousegen@har.mrc.ac.uk

**FEATURES**

source location/Qualifiers

1..197912  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /chromosome="9"  
 /clone="RP23-115A1"  
 /clone\_1fb="RPI-23"  
 /clone\_2fb="RPI-23"  
 BASE COUNT 54511 a 47510 c 46863 g 49028 t  
 ORIGIN

**Alignment Scores:**

Pred. No.: 3,61e+03 Length: 197912  
 Score: 89.50 Matches: 55  
 Percent Similarity: 41.56% Conservative: 41  
 Best Local Similarity: 23.81% Mismatches: 91  
 Query Match: 8.00% Indels: 46  
 DB: 10 Gaps: 11

US-10-034-500-2 (1-218) x AL672070 (1-197912)

QY 6 SetCythrlyarvalgluseryrAsnTyrluValaPyrSerGlySerMetMet 25  
 DB 83290 TCCTGTAAGCCCAAGAGGCGCTTAT-----CTGCAGGCA 83325  
 QY 26 MetLysHisValaValaValaGluProLysIleGluLeuAlaLysGlu-----Ala 42  
 DB 83326 CACACCCACCTCAACATCCCTGTAACCTTCACTCTGATCTGACACTTCTCT 83385  
 QY 43 IleLeuLysIleAlaAlaMetProLysMetSerTyrlngLysLeuTyrlthPhe 62  
 DB 83386 CAGATCAAGTACGACACTGTCTCACTCTCACACAGAGAGGCACTCGACTATC 83445  
 QY 63 Ala-----ProTySerValIlelleProLysIleSerTyrlpaSerCys 77  
 DB 83446 TGTCTGACAGCGCGCTTCTCTGACACCCCTGTAAC-AAATCAACAACCTCCCATGC 83504  
 QY 78 ValAlaGlu-----CysAlaValaLysThrIleLysSerAlaPheGluIlePheGlyArg 95  
 DB 83505 CCAGCGAGTCCCTCTCAAGCAGGCAAGGCACTCTGAGTCAAGGTAACACCTG 83564

QY 96 LeuThrProValGlyAspGlyIleLysMet-----HisGluThrValIleAsnGln 112  
 DB 83565 CTGTGTCACAG-GGAAGCGGCTGACTTTATCTTCCATCATCAGACAGGACAGT 83623  
 QY 113 MetProProGlnAlaIleValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsn 132  
 DB 83624 CAAGAAGTAAAGCTCTGAGTGTCTGTAACA-----CACACTCCACAGGCGCAAGC 83677  
 QY 133 ProValGluGluValLysSerIleTyrlngLysProAsnValCysPheHisValVal 152  
 DB 83678 CCGAGTCTCTTTTTCAGATTAACCCACAGAGCACTCCAAAC-----CCCATATC 83728  
 QY 153 Ser-----PheAlaAspAlaGluGlyLysAlaIle 163  
 DB 83729 TCCTAATACCTTCCCTCCACTGACACCAAGATTTCAGATTCGAAAGAGGAGGAG 83788  
 QY 164 IleAspGlnIleValaIleAlaLeuAsnSerGlySerValIleValaAspGlyLeuGlnLeu 182  
 DB 83789 AAAGATTAATAATTAACAAACAAACAAAGCCCTCAATGTTCTATTAATGAAATTAACA 83848  
 QY 183 ---LeuGlnAsnProAla---ValCysGlnGluPheValaLysSerValPheCysGlnGlu 200  
 DB 83849 AATTACCAAGAAACCAATTAACTTTGTAAG-----TATAACAAAGA 83890  
 QY 201 GlnIleLeuValThrGluGluValaValaValLeu 211  
 DB 83891 CAGATTTTATTTCTGAGCTCTCTGGAAGTTTG 83923

**RESULT** 40  
**LOCUS** AF224491 3429 bp mRNA linear INV 02-FEB-2001  
**DEFINITION** Halocynthia roretzi complement factor B mRNA, complete cds.  
**ACCESSION** AF224491  
**VERSION** AF224491.1 GI:12655865  
**KEYWORDS**  
**SOURCE** Halocynthia roretzi.  
**ORGANISM** Halocynthia roretzi.  
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;  
 Stolidobranchia; Pyrosidae; Halocynthia.

**REFERENCE**  
**AUTHORS** 1 (bases 1 to 3429)  
 Ji, X., Nankawa-Yamada, C., Nakanishi, M., Sasaki, M. and Nonaka, M.  
**TITLE** Unique domain structure of ascidian complement factor B: Trace of exon shuffling  
**JOURNAL** Unpublished  
**REFERENCE** 2 (bases 1 to 3429)  
 Ji, X., Nankawa-Yamada, C., Nakanishi, M., Sasaki, M. and Nonaka, M.  
**AUTHORS** Direct Submission  
**TITLE** Submitted (13-JAN-2000) Department of Biological Sciences, Graduate School of Science, University of Tokyo, 7-3-1 Hongo, Tokyo 113-0033, Japan

**FEATURES**

source location/Qualifiers

1..3429  
 /organism="Halocynthia roretzi"  
 /db\_xref="taxon:7729"  
 /cissue="type="hepatopancreas"  
 23..3277  
 /codon\_start=1  
 /product="complement factor B"  
 /protein\_id="AAK00631.1"  
 /db\_xref="GI:12655866"  
 /translation="MKLAVIFLCNVLLIPVYVYGVNSNSPFRFTCKDTCIPLVAKCD  
 GNPDGDEDELCTGSKAVKACATGALNGSLGYSRSCIPRIIFSGAGNACI  
 CDQEGYSCINPHSLCRGEVPLVENSYGHEIIRPDENVKRLPDKGDIYRYK  
 CRGRILVGLPRCKGKQWSENQOCVADRCGTPLVYRGVYSYTRNRRISGYP  
 IYGVKXYICPDHLLIGAPVCTKHRSBDRQCTVPCPRDSFAFVAGIAGPN  
 CITSACDGRLLICHDLCGRACVDNCTGAPTIQLAATVYTGSGOLARYTCE  
 GYVNLGAVLLCDSGSGKSLSTCIAMKGNPDRSLEFTSGVSLANRWYGSVAE  
 FSCPNRYMLGPTRVCMENMGSGGLTCDLKDNYILCPDGPVINGEIKGIFT  
 IGATVTFECHNGYVVGAEITCLYFKQMSNDPILCVDPAYSSKSDIARTLYKLSN  
 LPDGLTTSRTISASEVNI FHVIVITFVVSASVTKYDQFSGSLFAKRLIDRLKNG  
 GVLYKSIAYVASNKTGLEIDRFSTNKEVILKRLDNDSDVKEAVSLIBETRSGTA  
 TAKALKSRIDMMLFMEHDIRNDQTDKCHVFLFTDGMNEGNPVEVKEKQKIFGNS



IEFYISAOEDPSPFAFEELIGLASEPENVIYIEDIHLSSYLKTDVKSPKSCQ  
AGEVGLKRMALKRIVGSESVENAMPQALITKASTIKDIYTLGNOGGSLINDQ  
WVLTAHLFDRLKGBEDNMHESVLVHLGISIKTSEDMISIMYIIGBI IIPRYD  
KNTLNDVTLILGKEYHRMNTSYIERISYTPYIRPYCLPCNNISCLSKESQLTNDGK  
SLNGGDROCDIEERKILEENNAKYVAFGDTSRKNEPDPKKNIKLSKULQALIKI  
ODDSCODAIRIYINEMKORIKRYSYNTTLFPCIDPHDNGVDTCCGDSGGFVAREVLANOTS  
GVSCWVQIGLVSFGMCGALRDVHAIVPGFTYIVIKIIPWIKRKNMTDEAKL"  
BASE COUNT 1088 a 623 c 812 g 906 t  
ORIGIN

Alignment Scores:  
Pred. No.: 28.8 Length: 3429  
Score: 89.00 Matches: 46  
Percent Similarity: 37.27% Conservative: 36  
Best Local Similarity: 20.91% Mismatches: 70  
Query Match: 7.95% Indels: 68  
DB: 3 Gaps: 9

US-10-034-500-2 (1-218) x AF224491 (1-3429)

Qy 2 GluValThrAlaSerCysThrLysArgValGluSerTyrAsnTyrLeuValAspTyrSer 21  
Db 1628 GATGTTTCAAAAAGCGTGAACAAAGAAATACCAAGATTAGT-----AGT 1672  
Qy 22 GlySerMetMetLysHisValAlaArgGluProLysIleGluLeuAlaLysGlu 41  
Db 1673 GGG-----ATTGAATTGCGAAAGG 1693  
Qy 42 AlaIleLeuLysIleAsnAlaAlaMetProLysMetSerTyrGlnGlyGlyLeuTyrThr 61  
Db 1694 TTGATA-----GATAGGTGAAAATTTTGGTGGAGTATTGAAA 1732  
Qy 62 PheAlaProLysSerValIleIleProGlnGlySerTyrPheSerCysValAlaGluCys 81  
Db 1733 TATTCa-----ATCATTGCTTACGCATCATCAACAAACAAATTCAGAGATA 1780  
Qy 82 AlaValAsnThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGlyAsp 101  
Db 1781 ACAGATCCCTTTTCGACGAATGTAAAGAGTTATTAAAGTCTGATTAACCTGGATTGG 1840  
Qy 102 GlyIleLys----- 104  
Db 1841 CAGGTTAAAGAGCGTGTTCGAACTCATAGAGACAAAGAGTGTACGGCAACTGCA 1900  
Qy 105 -----MetHisGluThrValIleAsn 111  
Db 1901 AAAGCTTGAAGAAAGCTTCGAGACATGATGTTATTCATGAAACATGACATTGAATGAT 1960  
Qy 112 GlnMetProProGlnAlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMet 131  
Db 1961 CAGACAAATGACAAAGTGCATGTTCTTCTTACGACGAATGCACAAATGAGAGAAAG 2020  
Qy 132 AsnProValGlu-----GluValLysSerIleTyrGlnThrAsnProAsnValCys 148  
Db 2021 AATCTGTTGAAGTTCGAAAAGAAATGCAAAAGATATT-----GGCTTAATATAGAG 2074  
Qy 149 PheHisValIleSerPheAlaAspAspAlaGluGlyLysAlaIleIleAspGlnIleVal 168  
Db 2075 TTTTACAGTATTTCAGACACAAAGATCCAGTCCGAGGCA---TTGAGAGACTGATC 2131  
Qy 169 AlaLeuAsnSer-----GlySerValLeuValAspGlyLeuGlnLeuLeuGlnAsn 185  
Db 2132 GGACTTGCTTCTGAGCGTGAAGAAATTAATTATTCATCGAGGACATTCATCTTCAGTTCT 2191

Search completed: April 8, 2003, 05:46:57  
Job time : 2581 secs





GenCore version 5.1.4 p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 8, 2003, 03:50:09 ; Search time 215 Seconds  
(without alignments)  
2283.422 Million cell updates/sec

Title: US-10-034-500-2  
Sequence: 1 AEVYASCTKRVESYNLYVDY.....OEQILVTEEVVVLGVNFAF 216

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Delop 6.0, Delext 7.0  
Delop 6.0, Delext 7.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODL=frame+ p2n.model -DEV=xlh  
-O=/cgn2.1/USPRO.spool/US10034500/runat.02042003.092744.19692/app.query.fasta.1.391  
-DB=NGeneSeq.101002 -QFMT=fastcp -SUFFIX=ring -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=Blomum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=DCT -THR MAX=100 -THR MIN=0 -ALIGN=40  
-MODE=LOCAL -OUTFMT=prco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10034500.OCGN.1.1.200 @runat.02042003.092744.19692 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MMAP -LARGEQUERY -NEG SCORES=0 -MATT -LONGLOC -DEV TIMEOUT=120  
-MARK TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: /SID52/gcgcdata/geneSeq/geneSeq-emb1/NA1980.DAT.\*  
2: /SID52/gcgcdata/geneSeq/geneSeq-emb1/NA1981.DAT.\*  
3: /SID52/gcgcdata/geneSeq/geneSeq-emb1/NA1982.DAT.\*  
4: /SID52/gcgcdata/geneSeq/geneSeq-emb1/NA1983.DAT.\*  
5: /SID52/gcgcdata/geneSeq/geneSeq-emb1/NA1984.DAT.\*  
6: /SID52/gcgcdata/geneSeq/geneSeq-emb1/NA1985.DAT.\*  
7: /SID52/gcgcdata/geneSeq/geneSeq-emb1/NA1986.DAT.\*  
8: /SID52/gcgcdata/geneSeq/geneSeq-emb1/NA1987.DAT.\*  
9: /SID52/gcgcdata/geneSeq/geneSeq-emb1/NA1988.DAT.\*  
10: /SID52/gcgcdata/geneSeq/geneSeq-emb1/NA1989.DAT.\*  
11: /SID52/gcgcdata/geneSeq/geneSeq-emb1/NA1990.DAT.\*  
12: /SID52/gcgcdata/geneSeq/geneSeq-emb1/NA1991.DAT.\*  
13: /SID52/gcgcdata/geneSeq/geneSeq-emb1/NA1992.DAT.\*  
14: /SID52/gcgcdata/geneSeq/geneSeq-emb1/NA1993.DAT.\*  
15: /SID52/gcgcdata/geneSeq/geneSeq-emb1/NA1994.DAT.\*  
16: /SID52/gcgcdata/geneSeq/geneSeq-emb1/NA1995.DAT.\*  
17: /SID52/gcgcdata/geneSeq/geneSeq-emb1/NA1996.DAT.\*  
18: /SID52/gcgcdata/geneSeq/geneSeq-emb1/NA1997.DAT.\*  
19: /SID52/gcgcdata/geneSeq/geneSeq-emb1/NA1998.DAT.\*  
20: /SID52/gcgcdata/geneSeq/geneSeq-emb1/NA1999.DAT.\*  
21: /SID52/gcgcdata/geneSeq/geneSeq-emb1/NA2000.DAT.\*  
22: /SID52/gcgcdata/geneSeq/geneSeq-emb1/NA2001.DAT.\*  
23: /SID52/gcgcdata/geneSeq/geneSeq-emb1/NA2001B.DAT.\*  
24: /SID52/gcgcdata/geneSeq/geneSeq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	93.5	8.4	687	24 ABK75142	Bacillus licheniformis
2	92.2	8.2	2823	22 AAH98375	Human EST-derived
3	91.5	8.2	2616	21 AA64335	Clone 216888 of a
4	91.5	8.2	3043	24 AAD35923	Human C5193 full 1
5	91.5	8.2	3181	24 AAD5925	Human C5193 consen
6	91.5	8.2	3195	24 ABA91323	Human colon specif
7	91.5	8.2	3199	22 ABA09217	Human membrane-bou
8	91.5	8.2	3265	21 AA265095	Membrane-bound pro
9	91.5	8.2	3265	22 AA546053	Human DNA encoding
10	91.5	8.2	3265	22 AAP92092	Human PRO1124 CDNA
11	91.5	8.2	3265	22 AAP4421	Human PRO1124 (UNQ
12	90.5	8.1	37716	23 AAS59553	Protonibacterium
13	90	8.0	1380	22 AA171344	A. nidulans NADP-d
14	90	8.0	1433	22 AA171343	Aspergillus nidula
15	90	8.0	2944528	24 ABA03041	Listeria monocytog
16	87.5	7.8	2958	24 AB141969	DNA sequence of H4
17	87.5	7.8	2961	24 AB141970	DNA sequence of H4
18	87	7.8	1626	20 AAB61746	B. burgdorferi ant
19	87	7.8	1656	22 AAH00718	B. burgdorferi ant
20	87	7.8	1728	20 AA61745	Borrelia burgdorfe
21	87	7.8	910715	20 AA20248	DNA encoding novel
22	86.5	7.7	1404	23 AAS92979	DNA encoding novel
23	86	7.7	3667	21 AA52180	DNA encoding H. in
24	86	7.7	3685	21 AA52179	Haemophilus influe
25	86	7.7	88421	24 AAL40781	8842int genomic DN
26	86	7.7	3011208	24 AB069245	Listeria innocua D
27	85.5	7.6	3720	21 AA251000	X. oryzae avrXo1
28	85.5	7.6	6292	22 AAF60237	Human polynucleoti
29	85.5	7.6	6317	24 AAD38654	Human h217 secret
30	85.5	7.6	6373	22 AAF30188	Human 2143858 enc
31	85.5	7.6	6378	22 AAF15841	Human polynucleoti
32	84.5	7.6	2602	22 AAF1084	Human polynucleoti
33	84.5	7.6	6049	22 AAF19298	Human polynucleoti
34	84	7.5	3677	21 AA475149	CDNA encoding a hu
35	84	7.5	3677	21 AA475157	CDNA clone encodin
36	84	7.5	3789	13 AAQ50183	Bovine brain clone
37	84	7.5	52253	21 AAF21544	N. meningitidis pa
38	84	7.5	349980	21 AAF21544	Neisseria meningit
39	83.5	7.5	2573	19 AAV59153	Human ReproSA-1 co
40	83	7.4	2135	21 AAC39216	Arabidopsis thalia
41	83	7.4	3677	21 AA475158	CDNA clone encodin
42	83	7.4	3677	21 AA475159	CDNA clone encodin
43	83	7.4	3942	20 AA495550	Nucleic acid seque
44	82.5	7.4	816	21 AAF08436	Fusarium venenatum
45	82.5	7.4	2754	21 AAC49068	Arabidopsis thalia

## ALIGNMENTS

RESULT 1  
ID ABK75142 standard; DNA; 687 BP.  
ABK75142  
AC  
XX  
AC ABK75142;  
XX  
DT 13-AUG-2002 (first entry)  
XX  
DE Bacillus licheniformis genomic sequence tag (GST) #2433.  
XX  
KW Differential gene expression; genomic sequenced tag; GST;  
KW altered culture condition; environmental stress;  
KW physiological provocation; ds.  
XX  
OS Bacillus licheniformis.  
XX  
PN WO200229113-A2.  
XX

XX	11-APR-2002.
XX	
XX	05-OCT-2001; 2001WO-US31437.
XX	
XX	06-OCT-2000; 2000US-0680598.
XX	27-MAR-2001; 2001US-279526P.
XX	
XX	(NOVO ) NOVOZYMES BIOTECH INC.
XX	(NOVO ) NOVOZYMES AS.
XX	
XX	Berka R, Clausen IG;
XX	
XX	WPI; 2002-416684/44.
XX	
XX	Monitoring differential expression of several genes in first Bacillus
XX	cell relative to expression of same genes in one or more second
XX	Bacillus cells, by using substrate containing Bacillus genomic
XX	sequenced tag array -
XX	
XX	Claim 4; SEQ ID NO 2433; 200pp; English.
XX	
XX	The invention describes a method of monitoring differential expression of
XX	genes in a first Bacillus cell relative to expression of the genes in
XX	other Bacillus cells, comprising hybridising labelled nucleic acid probes
XX	isolated from Bacillus cells to a substrate containing array of Bacillus
XX	genomic sequenced tags (GST), examining the array, and determining
XX	relative gene expression by an observed hybridisation reporter signal of
XX	a spot in the array. The method is useful for measuring the expression of
XX	genes in a first Bacillus cell relative to expression of the same genes
XX	in one or more second Bacillus cells. The method is useful for monitoring
XX	global expression of several genes from a Bacillus cell, discovering new
XX	genes, identifying possible functions of unknown open reading frames and
XX	monitoring gene copy number variation and stability. Monitoring changes
XX	in expression of genes may be used to provide a representation of the way
XX	in which Bacillus cells adapt to changes in culture conditions,
XX	environmental stresses or other physiological provocation. Extensive
XX	follow-up characterisation is unnecessary, when one spot on an array
XX	equals one gene or one open reading frame, since sequence information is
XX	available. This sequence represents a genomic sequence tag (GST) used in
XX	the method of the invention.
XX	Note: The sequence data for this patent did not form part of the printed
XX	specification, but was obtained in electronic format directly from WIPO
XX	at
XX	ftp.wipo.int/pub/published_pct_sequences.
XX	
XX	Sequence 687 BP; 218 A; 147 C; 177 G; 145 T; 0 other:
XX	
XX	Alignment Scores:
XX	Pred. No.: 0.0725 Length: 687
XX	Score: 93.50 Matches: 46
XX	Percent Similarity: 34.43% Conservative: 27
XX	Best Local Similarity: 21.70% Mismatches: 84
XX	Query Match: 8.36% Indels: 55
XX	DB: 24 Gaps: 8
XX	
XX	US-10-034-500-2 (1-218) x ABK75142 (1-687)
XX	
XX	17 LeuValAspTrySerGlySerMetMetMetLeuHisValAlaValaGluProLysIle 36
XX	Db 124 ATGCTTGATCCAAAGCGGACAGTGAAGCTTAAAGATGCGGGGTC-----TCAAAATAC 177
XX	37 GluLeuAlaLysGluAlaIleLeuLysIle-----AsnAlaAla 49
XX	Db 178 GACCTTCCCAAAATGAAGCGTTACAGCTTGCTTCCAAAGCTTGAAAATGCAAAAGTGCTC 237
XX	50 MetProLysMetSerTyrGlnGly----- 57
XX	Db 238 ATGAGAGTTTTCGGTTCCAGAGGGAACAATAAATTCGGGAAAAGTCAGTCTGCAAT 297
XX	58 -----GlyLeuTyrTrpPheAlaProLysSerValIleIleProGlnGlySerTrp 74
XX	298 GCAATCAGAGGGTGGTACGGGCTTCCAAACGTATAC----- 333

QY 75 AenserCyValAlaGluCySalValAenThrIleYsSerAspLeuGluIlePheGly 94  
 Db 334 -----GAGCAAAAGCTTCGCCAATTCCTTAAACGGCATCGGA 369  
 QY 95 ArgLeu-----ThProValGluYAspGlyIleYsMethIsgluThrValIleAenGln 112  
 Db 370 CCGACCGGATGGACACCGCATCGCAAAACGGCTGCAGATGCAGAAACAGCGCTTGACCGAG 429  
 QY 113 MetProProGlnAla-----AlaValIleLeuLeuThrAspGlyYHsAsnAsnLeuGly 130  
 Db 430 CTGACACACAAACAGGGAAGAAACGTCGTCATCTGCTGCTACAGACGGTGAAGAAACATCGCGGA 489  
 QY 131 MetAsnProValGluGluValYsSerIleYrGlnThrAsnProAsnValCysPheHle 150  
 Db 490 GGCAATCCCGGTAAAGATCGCAACAGACAGTCGGCAAAATCCAAATCGCGTGGT-----AAC 543  
 QY 151 ValValSerPheAlaAspAspAlaGluGlyYrAlaIleIleAspGlnIleValAlaLeu 170  
 Db 544 GGCATCGCGCTTGATTTATATAGAGACCTTCCACGGCAATTCACAGTATCGCAGCAGCT 603  
 QY 171 AasnSerGlySerValLeuValAspGlyLeuGlnLeuGlnAsnProAlaValCysGln 190  
 Db 604 GCGCGCGGTGATATTTC-----CAGCGAAAACACTAA 636  
 QY 191 GluPheValAasnSerValPheCysGlnGlnGlnIle 202  
 Db 637 AATGACATCAAAAGAAATTTTACTCTCAGAGCAAT 672  
 RESULT 2  
 AAH98375  
 ID AAH98375 standard; cDNA; 2823 BP.  
 AC AAH98375;  
 XX  
 DT 12-OCT-2001 (first entry)  
 XX  
 DE Human EST-derived coding sequence SEQ ID NO: 232.  
 XX  
 KW Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;  
 KM tomato; monkey; dog; sea urchin; expressed sequence tag; EST;  
 KW diagnostic; forensic test; gene mapping; genetic disorder;  
 KM biodiversity; gene therapy; nutrition; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO200154477-A2.  
 XX  
 PD 02-AUG-2001.  
 XX  
 PF 25-JAN-2001; 2001WO-US02687.  
 XX  
 PR 25-JAN-2000; 2000US-0491404.  
 PR 17-JUL-2000; 2000US-0617746.  
 PR 03-AUG-2000; 2000US-0631451.  
 PR 15-SEP-2000; 2000US-063870.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;  
 PI Cao Y, Drmanac RA, Zhang J, Werhman T;  
 XX  
 DR WPI; 2001-476164/51.  
 DR P-PsDB; AAM23716.  
 PT Isolated polypeptide for treatment of diseases, diagnostics, raising  
 PT antibodies and research use -  
 XX  
 PS Claim 1; Page 341-342; 1275bp; English.  
 CC The present invention provides the protein and coding sequences of novel  
 CC proteins from a variety of organisms, including human, dog, cat, horse,  
 CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea  
 CC urchin and tomato. These were derived from expressed sequence tags (ESTs)

CC from the organism of interest. They can be used in diagnostics,  
CC forensics, gene mapping, identification of mutations, to assess  
CC biodiversity and for nutritional purposes. The present sequence is a cDNA  
CC of the invention.

**SQ** Sequence 2823 BP; 720 A; 744 C; 788 G; 571 T; 0 other

Alignment Scores:	
Pred. No.:	0.882
Score:	99.00
Percent Similarity:	46.72%
Best Local Similarity:	2.88%
Query Match:	8.22%
DB:	22
	7
Length:	2833
Matches:	34
Conservative:	30
Mismatches:	45
Indels:	28
Gaps:	7

```

Qy 16 TyLeuValAlaPrrYSerGlySerMetMetMetLeuValAlaValAlaArgLutProlys 35
Db 858 TTTGTGATTGACATCAAGGGCTCCAGTGGCTGTGGG-----AAA 896
Qy 36 lIleGluLeuAlaIysGluAlaIleLeuValIleAsnAlaAlaMetProIysMetSerTy 55
Db 897 TTAGAGCAGACAAAGAGGGCCCTTCTCAGAACTCTTGAAAGATATGAAAGGAACACT 956
Qy 56 GlmetGlyLeuTyThrPheAlaProTySerValIleIleProGlnGlySerTyrAsn 75
Db 957 -----CTGATTTTCATCTGTCTTCACTGAGAGATGTG-----TCCACATGGAAA 998
Qy 76 SerCyValAlaIaGlyVala-----ValAsnThrIleIysSerAsp 89
Db 999 GAGCACTTACTTCCAGGCCACGCCCGGAGAACCTTCCAGAGGCCAGGACGTTGTGAAGGC 1058
Qy 90 LeuGluIlePheGlyArgLeuThrProValGlyAspGlyIleIysMetHisGluThrVal 109
Db 1059 ATGAGAGATTAAAGA---ATGACCAACATTCAAAGAGGGCTGTGAGGGCACTCACTATG 1115
Qy 110 IleAsn-----GlnMetProProGlnAlaAla-----ValIleLeu 121
Db 1116 CTGAAACAAGGCCCGAGAGAGGACACAGATATCCAGAGAGAGGACCTCCATGTCTCATCATG 1175
Qy 122 LeuThrAspGlyHisAsnAsnLeuGlnIleAsnProValGluGluValLys 138
Db 1176 CTGACTATATGGAGATGCCCAATGTGTGTGAGAGAGACCCGAAAAAAATTCGA 1226

```

### RESULT 3

ID AAA64335 standard; DNA; 2616 BP.

AC AAA64335;

DT 20-DEC-2000 (first entry)

**DE** Clone 2516888 of a novel gene associated with colon disease.

KM Colonic disease; colon cancer gene; carbonic anhydrase; galactin;  
KM carcinoembryonic antigen; colorectal carcinoma tumour associated antigen  
KM fatty-acid binding protein; glutathione peroxidase; guanylin;  
KM cytokeratin; cadherin; intestinal mucin; colon cancer; Crohn's disease;  
KM metastatic colon cancer; atrophic gastritis; cholecystitis;  
KM irritable bowel syndrome; ulcerative colitis; ss.

**Homo sapiens**

PN WO200050588-A2.

PD 31-AUG-2000

PF 01-FEB-2000; 2000WO-US02595.

PR 22-FEB-1999; 99US-0255381.

PA (INCY-) INCYTE PHARM INC.

XX Walker MG, Volkmutz W, Klingler TM, Ial P;  
PI  
XX  
DR WPI; 2000-558397/51.

PT Novel gene associated with colon cancer and coexpressed with one or  
PT more known colon cancer genes in a number of biological samples, for  
PT use in gene therapy -  
XX  
PS Claim 2; Page 30-31; 33pp; English.

The present sequence represents a novel gene which is associated with colon disease. The gene is coexpressed with one or more colon cancer genes selected from carbonic anhydrase I, II and IV (CA I, II and IV), carcinoembryonic antigen family of proteins (cea), colorectal carcinoma tumour associated antigen (CO-029), down-regulated in adenoma (dra), fatty-acid binding protein (fabp), galectin (galec), glutathione peroxidase (gpx2), guanylin (guan), cytochrome 8 and 20 (ker 8 and 20), cadherin (cadher), or intestinal mucin (muc-2). The present polynucleotide sequence is useful for diagnosing a disease or condition associated with altered expression of the gene. The polynucleotide and encoded polypeptides are useful for treating or preventing such diseases. They are useful for diagnosing colon cancer, metastatic colon cancer, atrophic gastritis, cholecystitis, Crohn's disease, irritable bowel syndrome and ulcerative colitis.

**SQ** Sequence 2616 BP; 887 A; 488 C; 515 G; 726 T; 0 other;

Alignment Scores:		
Pred. No.:	0.91	Length: 2616
Score:	91.50	Matches: 46
Percent Similarity:	35.64%	Conservative: 26
Best local Similarity:	22.77%	Mismatches: 55
Query Match:	8.18%	Indels: 75

Qy	17	LeuValAspIyrSerGIySerMetMetMetLysHisValAlaValAGluProValIle	36
		:::::	
Db	388	GTTCCTTAAGTCGTGGAAAGCATGGG	414
Qy	37	GluDeuAlaIySGuValIleLeuLysIleAsnAlaIalMetProLysMetSerTyrGln	56
		:::::	
Db	415	-----GGTAAAGACCCGCTTAATCGAATGATCAAGACAGAAACATTTCCTGCTGACG	466
Qy	57	GlyGlyLeuTyrThrRheAlaProTyrSerValIleIleProGlnIySerTyrPAsnSer	76
		::::	
Db	469	-----ACTGTGAAATATGATCCTGCTGGGG	499
Qy	77	CysVal-----AlaGlu	80
Db	496	ATGGTTCACITTTGATGACTGCACATTTGTAATAGCTAATCCAAATTAAGACAGT	555
Qy	81	CysAlaValaenThrIleLysSerAspLeuGluIlePheGlyArgLysThrProValGly	100
		::::	
Db	556	GATGAAGAAACACACATCCATGGCAGGATTACCTACATAT-----CCTCTGGGA	603
Qy	101	-----AspGlyIleLysMetHisGluThrValIleAsnGlnMetProPro	115
		::::	
Db	604	GAAACTTCATCTGCTGCTGGAAATTAAATAGCATTTTCAGGTGATGGAGAGCATCATTC	666
Qy	116	Gln-----AlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsn	132
		::::	
Db	664	CAATCGATGGATCCGAAGTACTGCTGCTGATCTGATGGGAGGATACATCGCAAGTTCT	722
Qy	133	ProValGluGluValLysSerIleTyrGlnThrAsnProAsnValCysPheHisValVal	155
		::::	
Db	724	TGTATTGATGAAGTGAAGAA-----CAAAGTGGGGCATTTGATTTTATATGCTTTG	774
Qy	153	SerPheIa-----	155
Db	775	GGAAAGAGCTGCTGATGAGACGTAATAGATGAGACGAAGATTAACGAGAGAAATCATTTT	834

```

Oy 156 -----AspAspAlaGluGlyLysAlaIleIleAspGlnIleValAlaLeuSenser 172
Db 835 TATGTTTCAGATGAGAGCTGAGAACATGCGCTCATTTGCGGCTCTTACATCA 894
Oy 173 GlySer 174
Db 895 GGAAT 900

RESULT 4
AAD35923
ID AAD35923 standard; cDNA; 3043 BP.
AC AAD35923;
XX 26-JUL-2002 (first entry)
DE Human CS193 full length cDNA from clone 7741341H.
XX Human; CS193; gastrointestinal tract; cancer; gene therapy;
KW cytosolic; ss.
OS Homo sapiens.
XX US6368792-B1.
XX 09-APR-2002.
XX 27-MAR-1998; 98US-0049698.
XX PR 31-MAR-1997; 97US-0828856.
XX PA (ABBO ) ABBOTT LAB.
PI Billings PA, Cohen M, Colpits TL, Friedman PN, Hayden M;
PI Klaes MR, Roberts-rapp L, Russell JC, Stroupe SD;
XX WPI; 2002-328082/36.
XX DR WPI; 2002-328082/36.
XX PT New purified polynucleotide encoding CS193 antigen, useful for
XX PT diagnosing, staging, monitoring preventing or treating gastrointestinal
XX PT disorders -
XX PS Claim 1; Fig 1; 58bp; English.
XX CC The invention relates to a purified polynucleotide encoding CS193. The
XX CC polynucleotide is used for detecting, diagnosing, staging, monitoring,
XX CC prognosticating, preventing or treating diseases and conditions of the
XX CC gastrointestinal tract, particularly cancer. The CS193 gene is useful in
XX CC gene therapy. The present sequence is human CS193 full length cDNA.
XX CC Note: The present sequence is also shown in column 71-76 of the
XX CC specification, but lacks a nucleotide base at the end of each line.
XX SQ Sequence 3043 BP; 1023 A; 575 C; 608 G; 837 T; 0 other;

Alignment Scores:
Pred. No.: 1.14 Length: 3043
Score: 91.50 Matches: 46
Percent Similarity: 35.64% Conservative: 26
Best Local Similarity: 22.77% Mismatches: 55
Query Match: 8.18% Indels: 75
DB: 24 Gaps: 8

US-10-034-500-2 (1-218) x AAD35923 (1-3043)
Oy 17 LeuValAspTyrSerGlySerMetMetCysHisValAlaValArgGlnProGlySile 36
Db 941 GTTCTTGATTAAGTCTGGAACATGCGG----- 967
Oy 37 GluLeuAlaGlySerGluAlaIleLeuValIleAsnAlaAlaMetProGlyMetSerTyrGln 56
Db 968 -----GTTAAGACCGCTTAATCGAATGATCAACGACAAACATTCTCTGCTGACG 1021

```

```

Oy 57 GlyGlyLeuTyrThrPheAlaProTyrSerValIleIleProGlnGlySerTyrAsnSer 76
Db 1022 -----ACTGTTGAATAATGATCTCGGTGGCG 1048
Oy 77 CysVal-----AlaGlu 80
Db 1049 ATGTTCACTTGATGATGACTGCCACTATTGTAATTAAGCTAATCAATAAAGACAGT 1108
Oy 81 CysAlaValAsnThrIleYserAspLeuGluIlePheGlyArgLeuThrProValGly 100
Db 1109 GATGAAAGAAACACACTGACGAGATTACCTACATAT-----CCTCTGGGA 1156
Oy 101 -----AspGlyIleLysMetHisGluThrValIleAsnGlnMetProPro 115
Db 1157 GGAACCTTCATCTGCTGATTAATGATTAATGATTCAGTGATGAGAGCTACATTC 1216
Oy 116 Gln-----AlaAlaValIleLeuLeuThrAspGlyHisAsnAspLeuGlyMetAsn 132
Db 1217 CAACCTGATGATCCGAAGTACTGCTGACTGATGAGGAGATPACCTGCAAGTTCT 1276
Oy 133 ProValGluGluValIlySerIleTyrGlnThrAsnProAsnValCysPheHisValVal 152
Db 1277 TGTATTGATGAAGTAAA-----CAAAGTGGGCCATTGTTTATTTGCTTTG 1327
Oy 153 SerPheAla----- 155
Db 1328 GGAAGAGCTGCTGATGAGACGTAATAGATGACGACCAAGATPACGAGAGATCATTTT 1387
Oy 156 -----AspAspAlaGluGlyLysAlaIleIleAspGlnIleValAlaLeuSenser 172
Db 1388 TATGTTTCAGATGAGAGCTGAGAACATGCGCTCATTTGCGGCTCTTACATCA 1447
Oy 173 GlySer 174
Db 1448 GGAAT 1453

RESULT 5
AAD35925
ID AAD35925 standard; cDNA; 3181 BP.
AC AAD35925;
XX 26-JUL-2002 (first entry)
XX DT 26-JUL-2002 (first entry)
XX DE Human CS193 consensus sequence.
XX KW Human; CS193; gastrointestinal tract; cancer; gene therapy;
XX KW cytosolic; gene; ss.
XX OS Homo sapiens.
XX FT Key Location/Qualifiers
XX FT CDS 25..2778
XX FT /tag= a
XX FT /product= "Human CS193 polypeptide"

US6368792-B1.
PD 09-APR-2002.
PF 27-MAR-1998; 98US-0049698.
PR 31-MAR-1997; 97US-0828856.
XX PA (ABBO ) ABBOTT LAB.
XX PI Billings PA, Cohen M, Colpits TL, Friedman PN, Hayden M;
XX PI Klaes MR, Roberts-rapp L, Russell JC, Stroupe SD;
XX DR WPI; 2002-328082/36.
XX DR P-PSDB; AAE22687.
XX PT New purified polynucleotide encoding CS193 antigen, useful for

```

PT diagnosing, staging, monitoring preventing or treating gastrointestinal  
 PT disorders -  
 XX  
 PS Claim 1; Fig 1, 58pp; English.  
 XX  
 CC The invention relates to a purified polynucleotide encoding CS193. The  
 CC polynucleotide is used for detecting, diagnosing, staging, monitoring,  
 CC prognosticating, preventing or treating diseases and conditions of the  
 CC gastrointestinal tract, particularly cancer. The CS193 gene is useful in  
 CC gene therapy. The present sequence is human CS193 consensus sequence.  
 CC Note: The present sequence is also shown in column 75-80 of the  
 CC specification, but lacks a nucleotide base at the end of each line.  
 XX  
 SQ Sequence 3181 BP; 1080 A; 593 C; 631 G; 877 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 1 21 Length: 3181  
 Score: 91.50 Matches: 46  
 Percent Similarity: 35.64% Conservative: 26  
 Best Local Similarity: 22.77% Mismatches: 55  
 Query Match: 8.18% Indels: 75  
 DB: 24 Gaps: 8  
 US-10-034-500-2 (1-218) x AAD35925 (1-3181)  
 QY 17 leuValAspTyrSerGlySerMetMetMetLysHisValAlaValArgGluProLysIle 36  
 DB 952 GTTCTGTAAGTCTGGAAGCATGGCG----- 978  
 QY 37 GluLeuAlaLysGluAlaIleLeuLysIleAsnAlaIleMetProLysMetSerTyrGln 56  
 DB 979 -----GGTAAGACCCGCTAAATCGAATGATCAACAGCAAAACATTCTCTGTCGAG 1032  
 QY 57 GlyGlyLeuTyrThrPheAlaProTyrSerValIleIleProGlnGlySerThrAsnSer 76  
 DB 1033 -----ACTGTGAATATGATCCCTGGGTGGCG 1059  
 QY 77 CysVal-----AlaGlu 80  
 DB 1060 ATGGTTCACTTGATGACTGACCTATGTAATAAGCTAATCCAAATAAGCAAGT 1119  
 QY 81 CysAlaValAsnThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGly 100  
 DB 1120 GATGAAGAAACACACTCATGACGAGATTAACCTACATTT-----CTCTGGGA 1167  
 QY 101 -----AspGlyIleLysMetHisGluThrValIleAsnGluMetProPro 115  
 DB 1168 GGAACCTTCATCTGCTGCGAATTAATATGCAATTCAGGTGATGAGAGCTACATTCC 1227  
 QY 116 Gln-----AlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsn 132  
 DB 1228 CAATCCGATGATCGAAGTACTGCTGCTGACTGATGGGAGAGATACACGCAAGTCT 1287  
 QY 133 ProValGluGluValLysSerIleTyrGlnThrAsnProAsnValCysPheHisValVal 152  
 DB 1288 TGTATGTATGAAGTGA-----CAAGTGGGGCCATTGTCATTATTGCTTTG 1338  
 QY 153 SerPheAla----- 155  
 DB 1339 GGAAGAGCTGCTGATGAACAGTAATAGATGACAAAGATPAACAGAGAGTCAATTT 1398  
 QY 156 -----AepAepAlaGluGlyLysAlaIleIleAspGlnIleValAlaLeuAsnSer 172  
 DB 1399 TATGTTTCAGATGAAGCTCAGAAACAATGGCCTCATTGATGCTTTGGGGCTTTACATCA 1458  
 QY 173 GlySer 174  
 DB 1459 GGAAT 1464  
 RESULT 6  
 ID ABA91323 standard; DNA; 3195 BP.  
 XX

AC ABA91323;  
 XX  
 DT 08-APR-2002 (first entry)  
 XX  
 DE Human colon specific gene 370788.  
 XX  
 KW Colon specific gene; CSG; gene; human; colon cancer; tumour;  
 KW metastasis; diagnosis; gene therapy; imaging; staging; vaccine; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200192528-A2.  
 XX  
 PD 06-DEC-2001.  
 XX  
 PF 29-MAY-2001; 2001WO-US17583.  
 XX  
 PR 26-MAY-2000; 2000US-207383P.  
 XX  
 PA (DIAD-) DIADEXUS INC.  
 XX  
 PI Macina RA, Chen S, Plura J, Sun Y, Recipon H;  
 XX  
 DR WPI; 2002-083188/11.  
 XX  
 PT Colon specific polypeptides and polynucleotides useful for diagnosing,  
 PT monitoring, staging and treating colon cancer -  
 XX  
 PS Claim 1; Page 114-115; 116pp; English.  
 XX  
 CC The present sequence is that of a colon specific gene (CSG),  
 CC designated LSGold gene 370788. This is 1 of 22 claimed CSGs (see  
 CC ABA91302-23) identified in the Incyte LIFSEQ Gold database using  
 CC the data mining Cancer Leads Automatic Search Package (CLASP),  
 CC which identifies highly expressed organ and cancer specific genes.  
 CC The CSG, or its encoded protein, can be used as a diagnostic marker  
 CC of colon cancer. The invention provides claimed methods of  
 CC diagnosing colon cancer, monitoring colon cancer for the onset of  
 CC metastasis, and monitoring a change in stage of colon cancer in a  
 CC patient, by determining the level of a CSG in a sample of cells,  
 CC tissues or body fluid and comparing it with the level in a  
 CC corresponding sample from a healthy human control or from a sample  
 CC taken previously from the patient. Also claimed are methods  
 CC of identifying therapeutic agents for use in imaging and treating  
 CC colon cancer. Treatment may involve administering a compound  
 CC which downregulates expression or activity of a CSG. Adoptive  
 CC immunotherapy and gene therapy methods may be used. A claimed  
 CC vaccine for treating colon cancer comprises a claimed CSG.  
 XX  
 SQ Sequence 3195 BP; 1088 A; 596 C; 632 G; 879 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 1 22 Length: 3195  
 Score: 91.50 Matches: 46  
 Percent Similarity: 35.64% Conservative: 26  
 Best Local Similarity: 22.77% Mismatches: 55  
 Query Match: 8.18% Indels: 75  
 DB: 24 Gaps: 8  
 US-10-034-500-2 (1-218) x ABA91323 (1-3195)  
 QY 17 leuValAspTyrSerGlySerMetMetMetLysHisValAlaValArgGluProLysIle 36  
 DB 953 GTTCTGTAAGTCTGGAAGCATGGCG----- 979  
 QY 37 GluLeuAlaLysGluAlaIleLeuLysIleAsnAlaIleMetProLysMetSerTyrGln 56  
 DB 980 -----GGTAAGACCCGCTAAATCGAATGATCAACAGCAAAACATTCTCTGCGAG 1033  
 QY 57 GlyGlyLeuTyrThrPheAlaProTyrSerValIleIleProGlnGlySerThrAsnSer 76  
 DB 1034 -----ACTGTGAATATGATCCCTGGGTGGCG 1060









PF 28-FEB-2001; 2001WO-US06520.  
 XX  
 PR 01-MAR-2000; 2000WO-US05601.  
 PR 02-MAR-2000; 2000WO-US05841.  
 PR 03-MAR-2000; 2000US-187202P.  
 PR 06-MAR-2000; 2000US-186368B.  
 PR 14-MAR-2000; 2000US-189320P.  
 PR 14-MAR-2000; 2000US-189328P.  
 PR 15-MAR-2000; 2000WO-US06884.  
 PR 21-MAR-2000; 2000US-190828P.  
 PR 21-MAR-2000; 2000US-191007P.  
 PR 21-MAR-2000; 2000US-191048P.  
 PR 21-MAR-2000; 2000US-191314P.  
 PR 28-MAR-2000; 2000US-192655P.  
 PR 29-MAR-2000; 2000US-193032P.  
 PR 29-MAR-2000; 2000US-193053P.  
 PR 30-MAR-2000; 2000WO-US08439.  
 PR 04-APR-2000; 2000US-194449P.  
 PR 04-APR-2000; 2000US-194647P.  
 PR 11-APR-2000; 2000US-195975P.  
 PR 11-APR-2000; 2000US-196000P.  
 PR 11-APR-2000; 2000US-196187P.  
 PR 11-APR-2000; 2000US-196690P.  
 PR 11-APR-2000; 2000US-196820P.  
 PR 18-APR-2000; 2000US-198121P.  
 PR 18-APR-2000; 2000US-198585P.  
 PR 25-APR-2000; 2000US-199397P.  
 PR 25-APR-2000; 2000US-199550P.  
 PR 25-APR-2000; 2000US-199654P.  
 PR 03-MAY-2000; 2000US-201516P.  
 PR 17-MAY-2000; 2000WO-US13705.  
 PR 22-MAY-2000; 2000WO-US14042.  
 PR 30-JUN-2000; 2000WO-US14941.  
 PR 02-JUN-2000; 2000WO-US15264.  
 PR 05-JUN-2000; 2000US-209832P.  
 PR 28-JUL-2000; 2000WO-US20710.  
 PR 22-AUG-2000; 2000US-064484B.  
 PR 24-AUG-2000; 2000WO-US23328.  
 PR 08-NOV-2000; 2000WO-US30952.  
 PR 01-DEC-2000; 2000WO-US32678.  
 PR 20-DEC-2000; 2000WO-US34956.  
 XX  
 PA (GENTH) GENENTECH INC.  
 XX  
 PI Baker KP, Chen J, Deansoyers L, Goddard A, Godowski PJ, Gurney AL;  
 PI Pan J, Smith V, Deatanabe CK, Wood WI, Zhang Z;  
 XX  
 XX WPI; 2001-602746/68.  
 DR P-PSDB; AAU29152.  
 XX  
 PT Novel nucleic acids encoding PRO polypeptides, used to diagnose the  
 PT presence of tumours, such as prostate and breast tumours, in mammals and  
 PT to screen for modulators of the compounds -  
 XX  
 PS Claim 2; Fig 257; 774P; English.  
 XX  
 CC Sequences AAS45925-AAS46231 represent DNA molecules encoding and PCR  
 CC primers for PRO polypeptides of the invention. The sequences of the  
 CC invention can be used to detect the presence of a tumour in a mammal by  
 CC comparing the level of expression of a PRO polypeptide in a test sample  
 CC of cells from the animal and a control sample of normal cells, whereby a  
 CC higher level of expression in the test sample indicates the presence of a  
 CC tumour in the mammal. Mammals include dogs, cats, cattle, horses, sheep,  
 CC pigs, goats and rabbits but are preferably human. The polypeptides can be  
 CC used to stimulate tumour necrosis factor (TNF) alpha release from human  
 CC blood, when contacted with it. A specific polypeptide can be used to  
 CC stimulate the proliferation or differentiation of chondrocyte cells. The  
 CC PRO proteins can be used to determine the presence of tumours and also  
 CC susceptibility to tumour development, particularly adrenal, lung, colon,  
 CC breast, prostate, rectal, cervical, or liver tumours, in mammalian  
 CC subjects. The oligonucleotide probes specific for the PRO nucleic acids  
 CC can be used for genetic analysis of individuals with genetic disorders.  
 CC  
 XX

SQ Sequence 3265 BP; 1159 A; 596 C; 632 G; 878 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 1.26 Length: 3265  
 Score: 91.50 Matches: 46  
 Percent Similarity: 35.64% Conservative: 26  
 Best Local Similarity: 22.77% Mismatches: 55  
 Query Match: 8.18% Indels: 75  
 DB: 22 Gaps: 8  
 US-10-034-500-2 (1-218) x AAS46053 (1-3265)  
 Qy 17 LeuValaIepTyrSerGlySerMetMetLysHisValaIalaValArgGluProLysIle 36  
 Db 952 GTTCTGATAGTCTCGAAGCATGCGG----- 978  
 Qy 37 GluLeuAlaIysGluAlaIleLeuAlaIleLeuAlaIleMetProLysMetSerTyrGln 56  
 Db 979 -----GGTAAGACCCGCTTAATGAAATGAAATCAAGCAGAAACATTTCCTGCTGAG 1032  
 Qy 57 GlyGlyLeuTyrThrPheAlaProTyrSerValIleIleProGlnGlySerTyrPasnSer 76  
 Db 1033 -----ACTGTGAAATGATCCTGCGTGGG 1059  
 Qy 77 CysVal-----AlaGlu 80  
 Db 1060 ATGTTCACTTGGATGACTGACCATGTAATAAGTAAATCAATCAATAAAGCAGT 1119  
 Qy 81 CysAlaValaIanthrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGly 100  
 Db 1120 GATGAAGAAACACACTCATGCGAGGATTAACCTTACTT-----CCTCTGGGA 1167  
 Qy 101 -----AspGlyIleLysMetHisGluThrValIleAsnGlnMetProPro 115  
 Db 1168 GGAACCTTCATCTGCTCGGAATTAATATATGCAATTCAGGATTTGGAGACTACATTC 1227  
 Qy 116 Gln-----AlaAlaValaIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsn 132  
 Db 1228 CAATCGATGATCGAAGTACTGCTGCTGATGCGGAGATTAACATCGAAGTCT 1287  
 Qy 133 ProValGluGluValLysSerIleTyrGlnThrAsnProAsnValCysPheHisValVal 152  
 Db 1288 TGTATTGATGAAGTGAA-----CAAGTGGGCGCATTTGATTTATGCTTTG 1338  
 Qy 153 SerPheAla----- 155  
 Db 1339 GGAAGAGCTGCTGATGAAGCATTAATAGATGACGAAGATTAACAGGAAATCATTTT 1398  
 Qy 156 -----AspAspAlaGluGlyLysAlaIleIleAspGlnIleValaIleLeuAsnSer 172  
 Db 1399 TATGTTTCAGATGAAGCTCAGAACATGGCCTCATTTGCTTTGGGGCTCTTACATCA 1458  
 Qy 173 GlySer 174  
 Db 1459 GGAAT 1464  
 RESULT 10  
 ID AAF92092 standard; cDNA; 3265 BP.  
 XX  
 AC AAF92092;  
 XX  
 DT 15-MAY-2001 (first entry)  
 XX  
 DE Human PRO1124 cDNA.  
 XX  
 KW Human; PRO protein; mapping; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200116318-A2.  
 XX  
 PD 08-MAR-2001.

```

XX 24-AUG-2000; 2000WO-US23328.
PF 01-SEP-1999; 99WO-US20111.
XX 13-SEP-1999; 99WO-US21090.
PR 07-DEC-1999; 99US-0169495.
PR 09-DEC-1999; 99US-0170262.
PR 11-JAN-2000; 2000US-0175481.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 01-MAR-2000; 2000WO-US05601.
PR 03-MAR-2000; 2000US-0187202.
PR 25-APR-2000; 2000US-0199397.
PR 22-MAY-2000; 2000WO-US14042.
PR 05-JUN-2000; 2000US-0209832.
XX
PA (GENENTECH INC.
XX
PI Saton DL, Flivaroff E, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi CJ, Gurney AL, Matanabe CK, Wood WI,
XX
DR WPI; 2001-183260/18.
XX
DR P-PBDB; AAB87560.
XX
PT Eighty four nucleic acids encoding PRO polypeptides, useful in
PT molecular biology, including use as hybridization probes, and in
PT chromosome and gene mapping.
XX
PS Claim 2, Fig 69; 278pp; English.
XX
CC The present sequence is the coding sequence for a human PRO polypeptide
CC (secreted and transmembrane). The PRO protein, and PRO agonists, PRO
CC antagonists or anti-PRO antibodies are useful for preparation of a
CC medicament useful in the treatment of a condition which is responsive to
CC the PRO protein, agonists, antagonists or anti-PRO antibodies. The PRO
CC protein may also be employed as molecular weight markers for protein
CC electrophoresis. The PRO coding sequence has applications in molecular
CC biology, including use as hybridisation probes, and in chromosome and
CC gene mapping.
XX
SQ Sequence 3265 BP; 1159 A; 596 C; 632 G; 878 T; 0 other;
XX
Alignment Scores:
Pred. No.: 1.26 Length: 3265
Score: 91.50 Matches: 46
Percent Similarity: 35.64% Conservative: 26
Best Local Similarity: 22.77% Mismatches: 55
Query Match: 8.18% Indels: 75
DB: 22 Gaps: 8
US-10-034-500-2 (1-218) x AAF92092 (1-3265)
QY 17 LeuValAspTyrSerGlySerMetMetMetLeuHisValAlaValArgGluProLysIle 36
DB 952 GTTCTTGATGAAGCTGGAAGCATGGG----- 978
QY 37 GluLeuAlaLysGluAlaIleLeuLysIleAsnAlaAlaMetProLysMetSerTyrGln 56
DB 979 -----GTTAAGACCGCCCTAAATCGAATGAATCAAGACAAACAATTTCTGCTGCAG 1032
QY 57 GlyGlyLeuTyrThrPheAlaProTyrSerValIleIleProGlnGlySerTyrAsnSer 76
DB 1033 -----ACTGTAAATAATGATCTCTGGTGGG 1059
QY 77 CysVal-----AlaGlu 80
DB 1060 ATGTTCACTTGTATAGTACTGCACATATTGTAATAAGCTAATCAATAAAGAGAGT 1119
QY 81 CysAlaValAsnThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGly 100
DB 1120 GATGAAGAAGAACACACTCATGGCAGATTACTAACAATAT-----CCTCTGGGA 1167

```

```

QY 101 -----AspGlyIleLysMetHisGluThrValIleAsnGlnMetProPro 115
DB 1168 GGAATCTTCATCTGCTGATTAATAATATGATGATTCAGGTGAGAGCTACATTC 1227
QY 116 Gln-----AlaAlaValIleLeuLeuThrAspGlyHisAsnAspLeuGlyMetAsn 132
DB 1228 CAATCGATGATCCGAGGTCTGCTGCTGACTGATGAGGAGGATTAACATGCAATTTCT 1287
QY 133 ProValGluGluValLysSerIleTyrGlnThrAsnProAsnValCysPheHisValVal 152
DB 1288 TGTATTGATGAAGTGA-----CAAGTGGGGCCATTTTCATTTATTGCTTTG 1338
QY 153 SerPheAla----- 155
DB 1339 GGAAGAGCTGCTGATGAGACGATATAGATGACGACAAATTAACAGAGAAATCATTTT 1398
QY 156 -----AspAspAlaGluGlyLysAlaIleLeuAspGlnIleValAlaLeuAsnSer 172
DB 1399 TATGTTTCAGATGAGAGCTCAGAACATGCGCTCATTTGCTTTGGGGCTTTACATCA 1458
QY 173 GlySer 174
DB 1459 GGAAT 1464
RESULT 11
AAF44241
ID AAF44241 standard; cDNA; 3265 BP.
XX
AC AAF44241;
XX
XX 02-APR-2001 (first entry)
XX
DE Human PRO1124 (UNQ562) nucleotide sequence SEQ ID NO:378.
XX
KW Human; secreted and transmembrane protein; PRO; cytostatic;
KW cell death; cancer; chromosomal mapping; gene mapping; tissue typing;
KW diagnostic assay; ss.
XX
OS Homo sapiens.
XX
PN WO200073454-A1.
XX
PD 07-DEC-2000.
XX
XX 30-MAR-2000; 2000WO-US08439.
XX
PF 02-JUN-1999; 99WO-US12252.
PR 23-JUN-1999; 99US-0141037.
PR 07-JUL-1999; 99US-0143048.
PR 20-JUL-1999; 99US-0144758.
PR 26-JUL-1999; 99US-0145698.
PR 28-JUL-1999; 99US-0146222.
PR 17-AUG-1999; 99US-0149396.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 08-OCT-1999; 99US-0158663.
PR 30-NOV-1999; 99WO-US28313.
PR 01-DEC-1999; 99WO-US28301.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00376.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US04914.
PR 24-FEB-2000; 2000WO-US05004.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 20-MAR-2000; 2000WO-US07377.
XX
PA (GENENTECH INC.
XX

```



CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 37716 BP; 7884 A; 11879 C; 10798 G; 7150 T; 5 other;

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:
Score:	60.3	36	23
Percent Similarity:	90.50	23	52
Best Local Similarity:	44.70%	27.27%	21
Query Match:	8.09%	Indels:	7
DB:	23	Gaps:	7

US-10-034-500-2 (1-218) x AAS59553 (1-37716)

QY 18 ValAspTyrSerGlySerMetMetLysHisValAlaValArgGluPro---LysIle 36  
 DB 8611 ATTGACGTGTGGCTCATGTGGCGACGACGTT-----GACCCCTTCGACTT 8561  
 QY 37 GluLeuAlaLysGluAlaIleLeuLysIleAsnAlaAlaMetProLysMetSerTyrGln 56  
 DB 8560 TCCGCTGCTTAGACTGCTGCGAAGATTTCTGGGGGATTGCGGCA---CGTTCAAT 8504  
 QY 57 GlyGlyLeuTyrThrPheAlaProTyrSer---ValIleIleProGlnGlySerTyrPasn 75  
 DB 8503 GTCTCTTGTTGTAAGTTCGACGATCTGCTCAAGTGTGGTACCC-----8459  
 QY 76 SerCysValAlaGluCysAlaValAsnThrIleLysSerAspLeuGluIlePheGlyArg 95  
 DB 8458 CCCACACGCGATGTGCGCGGTATGACTGCGATCCAGACCTCCAGGCTGCCATCT 8399  
 QY 96 LeuThrProValGlyAspGlyIleLysMetHisGluThrValIleAsnGlnMetPro--- 114  
 DB 8398 ---ACCGCGATCGCGCGAGGGGATTATTCCTCCCTCAACGCTCAAACTGTCCGAC 8342  
 QY 115 -----ProGlnAlaAlaValIleLeuLeuThrAspGlyHis 126  
 DB 8341 GATCCGAACATCCGCGACGAGCAAGCCACGCGCGATCGTCTTCTTCTGACGGTGA 8282  
 QY 127 AsnAsnLeuGlyMetAsnProValGluGluValLys 138  
 DB 8281 ACGAATGTTGGCCGCTCTTCTTGAAGCGCCAG 8246

RESULT 13  
 AAI71344  
 ID AAI71344 standard; DNA; 1380 BP.  
 XX  
 AC AAI71344;  
 XX  
 DT 15-JAN-2002 (first entry)  
 XX  
 DE A. nidulans NADP-dependent glutamate dehydrogenase coding sequence.  
 XX  
 KW NADP-dependent glutamate dehydrogenase; GDH; AN-gdh-17; overexpression;  
 KM amino acid content; nutritional value; transgenic plant; tomato;  
 KW Lycopersicon esculentum; ds.  
 XX  
 OS Aspergillus nidulans.  
 XX  
 FH Key Location/Qualifiers  
 FT exon 1..322  
 FT /\*tag= a  
 FT /number= 1  
 FT misc\_feature 322..323  
 FT /\*tag= b  
 FT /note= "Splice junction between exons 1 and 2"  
 FT exon 323..1380  
 FT /\*tag= c  
 FT /number= 2  
 XX  
 XX JP2001238556-A.  
 XX  
 XX 04-SEP-2001.  
 PD

PF 18-DEC-2000; 2000JP-0404322.  
 XX  
 PR 16-DEC-1999; 99JP-0376710.  
 XX  
 PA (AJIN) AJINOMOTO KK.  
 XX  
 DR WPI; 2001-610051/70.  
 XX  
 PT Creation of a transgenic plant of improved amino acid composition for  
 PT use as a food, comprising culturing a plant transformed by a genetic  
 PT structure giving an excessive expression of glutamate dehydrogenase -  
 XX  
 PS Example 1; Fig 1-2; 47pp; Japanese.  
 XX  
 CC This sequence represents the coding sequence of the Aspergillus nidulans  
 CC NADP-dependent glutamate dehydrogenase (GDH) gene, referred to as  
 CC AN-gdh-17 in the specification. The invention relates to the generation  
 CC of transgenic plants (e.g., tomato) which overexpress GDH, the transgenic  
 CC plants themselves, and the seeds and progeny of such plants. The method  
 CC of the invention involves transforming a plant with a genetic construct  
 CC comprising a GDH sequence, a promoter such as the CMV (cauliflower  
 CC mosaic virus) 35S promoter or the tomato fruit-specific promoter 2A11,  
 CC and a marker gene, and then selecting plants on the basis of the  
 CC phenotype conferred by expression of the marker gene. The transgenic  
 CC plants of the invention accumulate larger quantities of free amino acids  
 CC in the edible tissues (e.g., fruit) relative to wild-type plants, giving  
 CC such plants a higher nutritional value.  
 XX  
 SQ Sequence 1380 BP; 256 A; 440 C; 361 G; 323 T; 0 other;

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:
Score:	0.548	1380	45
Percent Similarity:	90.00	25	60
Best Local Similarity:	37.63%	24.19%	56
Query Match:	8.04%	Indels:	8
DB:	22	Gaps:	8

US-10-034-500-2 (1-218) x AAI71344 (1-1380)

QY 17 LeuValAspTyrSerGlySerMetMetLysHisValAlaValArgGluProLysIle 36  
 DB 748 CTTTCCGACTCCAGAGGCTCTCTCATGTGCAAGATGATCGCTTTTCAACCCCTGA 807  
 QY 37 GluLeuAlaLysGluAlaIleLeuLysIleAsn-----AlaAlaMetProLys 52  
 DB 808 GAGATCGCCCTTATGCGCACTCAAGGTGCCCGCAACAACTCCGAGCTGCCACC 867  
 QY 53 MetSerTyrGlnGlyLeuTyrThrPheAlaPro-----64  
 DB 868 TCTCCGCTTGGCGCGAAGTTCACTATCGCATGCTCGCCCTTGACCAACATT 927  
 QY 65 -----TyrSerValIleIleProGlnGlySerTyrPasnSerCysValAlaGluCys 81  
 DB 928 CCGCGCACTTCAGAGTGTCTCTCTTCTGCACTGAGAAAGAGTCTCCGCGAAGAA 987  
 QY 82 AlaValAsnThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGlyAsp 101  
 DB 988 GCGAGACACCTCATCAAGTCCGCTCCGCTAT-----ATTGCTGAG 1029  
 QY 102 GlyIleLysMetHisGluThrValIleAsnGlnMetProProGlnAlaAlaValIleLeu 121  
 DB 1030 GGTTCACATGGGTGGACAC-----CAGGCGCCATCGACATC 1068  
 QY 122 LeuThrAspGlyHisAsnAsnLeuGlyMetAsnProValGluGluValLysSerIleTyr 141  
 DB 1069 TTT---GAGGCTCACCGCAAC-----GCCAACCCCGGAGATCGCATCGTAC----- 1113  
 QY 142 GlnThrAsnProAsnValCysPheHisValSerPheAlaAspAspAlaGluLys 161  
 DB 1114 -----GCCCTGTATA 1125  
 QY 162 AlaIleIleAspGlnIleValAlaLeuAsnSerGlySerValIleuValAspGlyLeuGln 181

```

Db 1126 -----GCCGCCAACGCTGGTGGTGGCCGCTCTCTGTGAG 1164
Oy 182 LeuLeuGlnAenProAla 187
Db 1165 ATGCTCAGAACTCTGCT 1182

RESULT 14
AA171343 standard; DNA; 1433 BP.
AC AA171343;
XX
XX
XX 15-JAN-2002 (first entry)
DE Aspergillus nidulans NADP-dependent glutamate dehydrogenase gene.
XX
XX NADP-dependent glutamate dehydrogenase; GDH; AN-gdh-17; overexpression;
KM amino acid content; nutritional value; transgenic plant; tomato;
XX Lycopersicon esculentum; ds.
XX
XX Aspergillus nidulans.
OS
XX
XX Key Location/Qualifiers
FT exon 1..322
FT /*tag= a
FT /number= 1
FT intron 323..375
FT /*tag= b
FT /number= 1
FT exon 376..1433
FT /*tag= c
FT /number= 2
XX
XX JP2001238556-A.
XX
XX 04-SEP-2001.
XX
XX 18-DEC-2000; 2000JP-0404322.
XX
XX 16-DEC-1999; 99JP-0376710.
XX
XX (AJIN ) AJINOMOTO KK.
XX
XX MPI, 2001-610051/70.
XX
XX Creation of a transgenic plant of improved amino acid composition for
PT use as a food, comprises culturing a plant transformed by a genetic
PT structure giving an excessive expression of glutamate dehydrogenase -
XX
XX Claim 4; Fig 1-2; 47pp; Japanese.
XX
XX This sequence represents the Aspergillus nidulans NADP-dependent
CC glutamate dehydrogenase (GDH) gene, referred to as AN-gdh-17 in the
CC specification. The invention relates to the generation of transgenic
CC plants (e.g., tomato) which overexpress GDH, the transgenic plants
CC themselves, and the seeds and progeny of such plants. The method
CC of the invention involves transforming a plant with a genetic construct
CC comprising a GDH sequence, a promoter such as the CMV (cauliflower
CC mosaic virus) 35S promoter or the tomato fruit-specific promoter 2A11,
CC and a marker gene, and then selecting plants on the basis of the
CC phenotype conferred by expression of the marker gene. The transgenic
CC plants of the invention accumulate larger quantities of free amino acids
CC in the edible tissues (e.g., fruit) relative to wild-type plants, giving
CC such plants a higher nutritional value.
XX
XX Sequence 1433 BP; 269 A; 448 C; 374 G; 342 T; 0 other;
SQ

```

```

Query Match: 8.04% Indels: 56
DB: 22 Gaps: 8
US-10-034-500-2 (1-218) x AA171343 (1-1433)
Oy 17 LeuValAenProAlaSerGlySerMetMetMetValAlaValArgGluProLysIle 36
Db 801 CTTTCCGACTCCAGAGGCTCTCATTTGTCAAGAGTAGTCCGCTTTTCAACCCCTGAA 860
Oy 37 GluMetAlaLysGluAlaIleLeuLysIleAsn-----AlaAlaMetProLys 52
Db 861 GAGATCGCCCTCATTTGCTCCGACTCAAGGTGCCGCAAGCAACTCTCCGAGCTCGCACCC 920
Oy 53 MetSerTyrgInglYglYLeuTyThrPheAlaPro----- 64
Db 921 TCCTCCGCTTTCGCGCAAGTTCACCTACATCCCATGCTCGCCCTTGACCAACATT 980
Oy 65 -----TyrservalIleIleProGlnIlySerTyraPanserCyvalAlaGluCys 81
Db 981 CCCGCAAGTTCGAGGTTGCTCTCCCTTCTGCACCTCAGAACGAAGTCCGCGAGGAA 1040
Oy 82 AlaValAenThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGlyAsp 101
Db 1041 GCCGAGCACTCATCACTCAAGTCCGCTGTCGGCTAT-----ATTGCTGAG 1082
Oy 102 GlyIleLysMetH1sgLutThrValIleAsnGlnMetProProGlnAlaAlaValIleLeu 121
Db 1083 GGTTCCAACATGGGTGTGAC-----CAGCGCCCATGCACATC 1121
Oy 122 LeuThrAspGlyH1sAsnAsnLeuGlyMetAsnProValGluGluValLysSerIleTy 141
Db 1122 TTT---GAGGCTCACCGCAC-----GCCAACCCCGCGATGCATGTGTAC----- 1166
Oy 142 GlnThrAsnProAsnValCysPheHisValSerPheAlaAspAspAlaGluGlyLys 161
Db 1167 -----GCCCCGTGTA 1178
Oy 162 AlaIleIleAspGlnIleValAlaLeuAsnSerGlySerValLeuValAspGlyLeuGln 181
Db 1179 -----GCCGCCAACGCTGGTGGTGGCCGCTCTCTGTGAG 1217
Oy 182 LeuLeuGlnAenProAla 187
Db 1218 ATGCTCAGAACTCTGCT 1235

RESULT 15
ABA03041/C
ID ABA03041 standard; DNA; 2944528 BP.
XX
XX ABA03041;
XX
XX 05-FEB-2002 (first entry)
XX
XX Listeria monocytogenes EGD-e genome sequence.
DE
XX Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KM vitamin B12; bacterial infection; disease; ds.
XX
XX Listeria monocytogenes.
OS
XX WO200177335-A2.
XX
XX 18-OCT-2001.
XX
XX 11-APR-2001; 2001WO-FR01118.
XX
XX 11-APR-2000; 2000FR-0004629.
XX
XX (INSP ) INST PASTEUR.
XX
XX Buchrieser C, Frangeul L, Couve E, Rusnok C, Fehli H, Dehoux P,
PI Duesbarger O, Chetouani F, Nedjari H, Glaser P, Kunst P, Cossart P,
PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;

```

PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;  
 PI Charrabory T, Doman R, Hain T, Berche P, Chablit A, Durant L;  
 PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;  
 PI Madueno E, de Pablo B, Wehland J, Kaerret U, Entian K, Hauf J;  
 PI Rose M, Voss H;  
 DR WPI; 2002-010914/01.  
 PT Genomic sequence for *Listeria monocytogenes*, useful e.g. for treatment  
 PT and prevention of *Listeria* and related bacterial infections, and  
 PT related polypeptides  
 PS Claim 1; SEQ ID No 1; 192bp; French.  
 CC The present sequence is the genome sequence of *Listeria monocytogenes*  
 CC EGD-e. This sequence and fragments of this sequence are useful for  
 CC selecting probes and primers for detecting genes in *L. monocytogenes* and  
 CC related organisms, and to study genetic polymorphisms and other genomes.  
 CC proteins (AB847297-AB850149) expressed from the present sequence are  
 CC useful for raising specific antibodies, identification of *L.*  
 CC *monocytogenes* and related organisms, and for biosynthesis and  
 CC biodegradation, especially biosynthesis of Vitamin B12. This sequence and  
 CC proteins encoded by it are also useful for selecting compounds that  
 CC regulate gene expression and cell replication and modulate *L.*  
 CC *monocytogenes*-related diseases. In addition, this sequence and proteins  
 CC encoded by it are useful in pharmaceutical and vaccine compositions for  
 CC the treatment or prevention of infections by *L. monocytogenes* and related  
 CC organisms.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 XX  
 SQ Sequence 2944528 BP; 914202 A; 563301 C; 555061 G; 911964 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 4.11e+04 Length: 2944528  
 Score: 90.00 Matches: 56  
 Percent Similarity: 36.47% Conservative: 37  
 Best Local Similarity: 21.96% Mismatches: 74  
 Query Match: 8.04% Indels: 88  
 DB: 24 Gaps: 15  
 US-10-034-500-2 (1-218) x ABA03041 (1-2944528)  
 QY 21 SerGlySerMetMetMetLysHisValAlaValArgLysProLysIleGluLeuAlaLys 40  
 DB 2516361 AACGGAAATCAATGATGCGCCCAATGCGAATGATTAATCCGAA----- 2516317  
 QY 41 GluAlaIleLeuLysIleAsnAlaIleMetProLysMetSerTyrGln---GlyGlyLeu 59  
 DB 2516316 -----GATGTCATGACGACGTGATTGATGACGATATATGTTGGTGACGAT 2516269  
 QY 60 TyrThrPheAlaProTyr-----SerValIleIleProGln 71  
 DB 2516268 TTAATTGTGTGACCAATCGTTCAAGAGGCCAAACAGAAAAGAGTTATTATTCACAGAA 2516209  
 QY 72 GlySerTyrPanserCysValAlaGluCysAlaValAsnThrIleLysSerAspLeuGlu 91  
 DB 2516208 GGGGAATGG-----GTTGAT 2516194  
 QY 92 IlePhe---GlyArgLeuThrProValGlyAspGlyIleLysMetHisGluThrValIle 110  
 DB 2516193 ATCTGGAACGGTGGGTGTCACCTGTCGGAACAATTTCTTATTATGCGCATGTA--- 2516137  
 QY 111 AsnGlnMetPro-----ProGlnAlaIleValIleLeu-----LeuThrAspGly 125  
 DB 2516136 GATACCGCTTCTGATTATTCGAAAGCGAGCAATATTCGATGATATGACAGATGGA 2516077  
 QY 126 HisAsnAsnLeuGlyMetAsnProValGluGluValLysSerIleTyrGlnThrAsnPro 145  
 DB 2516076 TAC---CAACTGTGTCAAAACGTGGAAACGATTAAATCCAT-----GAT 2516032  
 QY 146 AsnValCysPheHisVal-----ValSerPheAlaAspAspAla 158

DB 2516031 AATTAACCTTCCGAGTTTATCCATCTGAGATAGTAGTATAGCTTCTATGACGATGTA 2515972  
 QY 159 GluGly-----LysAlaIle 163  
 DB 2515971 AACGGTGGAGAAATGCGTACATCATGATGTTTTCGACACTTTGCCAAATGAAAAGTTCC 2515912  
 QY 164 IleAsp-----GlnIleValAlaLeuAsnSerGly 173  
 DB 2515911 GTACATCTTCCAGCATGGCAGATGAAACAACATGCAAGTTTCAGTACGAAACCAACA 2515852  
 QY 174 SerValLeuValAspGlyLeuGlnLeuGln-----AsnPro 186  
 DB 2515851 AGTGTACAAATTTGACGGGCGAGATGCTAAAGCAGATACATTAGATGCAATTTAACGAA 2515792  
 QY 187 AlaValCysGlnGluPheValAsnSerVal-----PheCysGlnGluGlnIle 202  
 DB 2515791 GCACCACTGGTTCTACTATGATATGCTACAAACTTAATATGTAAGAAACGACAGCA 2515732  
 QY 203 LeuValThrGluGluValAlaValAlaLeuArgGlyValAsnPheAla 217  
 DB 2515731 AAAGATGCCAAACAAGCATGCTAATGCTGTAACCATGCA 2515687  
 RESULT 16  
 ABL41969  
 ID ABL41969 standard; DNA; 2958 BP.  
 XX  
 AC ABL41969;  
 XX  
 DT 11-JUN-2002 (first entry)  
 XX  
 DE DNA sequence of H4P heavy chain of inter-alpha-inhibitor protein.  
 KW MAY-1; blood transfusion; immunosuppressive; mixed lymphocyte reaction;  
 KW MLR; autoimmune disease; allergic disease; organ rejection;  
 KW organ transplantation; rheumatism; psoriasis; bronchial asthma;  
 KW allergic rhinitis; allergic dermatitis; poliomyelitis; gene;  
 KW H4P heavy chain; inter-alpha-inhibitor protein; ss.  
 XX  
 OS Rattus sp.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 16..2814  
 FT /\*tag= a  
 FT /product= "H4P heavy chain of inter-alpha-inhibitor  
 FT protein"  
 FT  
 XX  
 PN WO200212495-A1.  
 XX  
 PD 14-FEB-2002.  
 XX  
 PP 01-AUG-2001; 2001WO-JP06620.  
 XX  
 PR 09-AUG-2000; 2000JP-0241169.  
 XX  
 PA (MARU-) MARUHO KK.  
 XX  
 PI Uchida H, Tanaka H, Kitoh Y, Fujimura A, Kobayashi E;  
 DR WPI; 2002-217191/27.  
 DR P-PSDB; ABB09706.  
 XX  
 PT Rat or human protein MAY-1 induced by homogeneous blood transfusion and  
 PT exhibiting immunosuppressive activity in a homogeneous mixed lymphocyte  
 PT reaction, for providing drug compositions to treat e.g. autoimmune  
 PT diseases  
 PS Disclosure; Page 47-49; 85pp; Japanese.  
 CC The present sequence encodes the rat H4P heavy chain of  
 CC inter-alpha-inhibitor protein. The specification describes MAY-1  
 CC protein, which is induced by homogeneous blood transfusion. MAY-1  
 CC exhibits an immunosuppressive activity in a homogeneous mixed



CC lymphocyte reaction (MDR). The MAY-1 protein can formulated into drug  
compositions for the prevention or treatment of autoimmune diseases,  
CC allergic diseases, or rejection reaction during organ transplantation,  
CC e.g. rheumatism, psoriasis, bronchial asthma, allergic rhinitis,  
CC allergic dermatitis and pollinosis.

XX Sequence 2958 BP; 774 A; 773 C; 759 G; 652 T; 0 other;

Alignment Scores:

Pred. No.: 3.42 Length: 2958  
Score: 87.50 Matches: 33  
Percent Similarity: 42.95% Conservative: 31  
Best Local Similarity: 22.15% Mismatches: 40  
Query Match: 7.82% Indels: 45  
DB: 24 Gaps: 9

US-10-034-500-2 (1-218) x ABL41969 (1-2958)

Qy 16 TyrLeuValAspTyrSerGlySerMetMetMetCysHisValAlaValArgGluProLys 35  
Db 841 TTTGTCATTGATTAAGCGGATCTATGGCAGGCAAG-----AAA 879  
Qy 36 IleGluLeuAlaLysGluAlaIleLeuLysIleAsnAlaAlaMetProLysMetSerTyr 55  
Db 880 ATCCACAGACCCGAGAGCCCTAATCAAGATC-----TTGAAAGACCTCAGACCC 930  
Qy 56 GlnGlyGlyLeuTyrThrPheAlaProTyrSerValIleIleProGlnGly----- 72  
Db 931 CAAGAC-----CAGTTCAATATCATTTGTTCACTGCGGAGCAAC 972  
Qy 73 SerTPaenserCyValAlaGluCysAlaValAsnThrIleLysSerAspLeuGluIle 92  
Db 973 CAGTGGAGCAGCTGCTGTCGACACACAGAGAACTTGAACAGCGCGTTGACTAT 1032  
Qy 93 PheGlyAspLeuThrProValGlyAspGlyIleLysMetHisGlnThrValIle----- 110  
Db 1033 GCTTCCAAGATC---CCGGCTCAGGAGGAGCAACATCAATAAGACAGTCTATCGGCT 1089  
Qy 111 -----AsnGln-----MetProGlnAlaAla----- 118  
Db 1090 GTGGAACGCTGATTAAGCAACAGCGCTAGCTAGCTCCCAAGAGCGTTCCCTC 1149  
Qy 119 ValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsnProValGluGluValLys 138  
Db 1150 ATCATCTGCTCAGCGATGGCAGGCCCACTGTGGG----- 1185  
Qy 139 SerIleTyrGlnThrAsnProAsnVal 147  
Db 1186 -----GAGACCAATCCCAAGATT 1203

RESULT 17

ABL41970 standard; DNA; 2961 BP.

ABL41970;

11-UN-2002 (first entry)

DNA sequence of H4P heavy chain of inter-alpha-inhibitor protein.

MAY-1; blood transfusion; immunosuppressive; mixed lymphocyte reaction;  
MLR; autoimmune disease; allergic disease; organ rejection;  
organ transplantation; rheumatism; psoriasis; bronchial asthma;  
allergic rhinitis; allergic dermatitis; pollinosis gene;  
H4P heavy chain; inter-alpha-inhibitor protein; ss.

Rattus sp.

Location/Qualifiers  
16.2817  
a

Key CDS  
/tag= a  
/trans\_except= (pos:301..304,aa:Ieu)  
/trans\_except= (pos:331..333,aa:Tyr)

/product= "H4P heavy chain of inter-alpha-inhibitor  
protein"

WO200212495-A1.

14-FEB-2002.

01-AUG-2001; 2001WO-JP06620.

09-AUG-2000; 2000JP-0241169.

(MARU-) MARUHO KK.

Uchida H, Tanaka H, Kitch Y, Fujimura A, Kobayashi E;

WPI: 2002-217191/27.

P-PSDB; ABB09707.

Rat or human protein MAY-1 induced by homogeneous blood transfusion and  
exhibiting immunosuppressive activity in a homogeneous mixed lymphocyte  
reaction, for providing drug compositions to treat e.g. autoimmune  
diseases -

Claim 7; Page 53-55; 85pp; Japanese.

The present sequence encodes the rat H4P heavy chain of  
inter-alpha-inhibitor protein. The specification describes MAY-1  
protein, which is induced by homogeneous blood transfusion. MAY-1  
exhibits an immunosuppressive activity in a homogeneous mixed  
lymphocyte reaction (MDR). The MAY-1 protein can formulated into drug  
compositions for the prevention or treatment of autoimmune diseases,  
CC allergic diseases, or rejection reaction during organ transplantation,  
CC e.g. rheumatism, psoriasis, bronchial asthma, allergic rhinitis,  
CC allergic dermatitis and pollinosis.

Sequence 2961 BP; 774 A; 773 C; 762 G; 652 T; 0 other;

Alignment Scores:

Pred. No.: 3.43 Length: 2961  
Score: 87.50 Matches: 33  
Percent Similarity: 42.95% Conservative: 31  
Best Local Similarity: 22.15% Mismatches: 40  
Query Match: 7.82% Indels: 45  
DB: 24 Gaps: 9

US-10-034-500-2 (1-218) x ABL41970 (1-2961)

Qy 16 TyrLeuValAspTyrSerGlySerMetMetMetCysHisValAlaValArgGluProLys 35  
Db 844 TTTGTCATTGATTAAGCGGATCTATGGCAGGCAAG-----AAA 882  
Qy 36 IleGluLeuAlaLysGluAlaIleLeuLysIleAsnAlaAlaMetProLysMetSerTyr 55  
Db 883 ATCCACAGACCCGAGAGCCCTAATCAAGATC-----TTGAAAGACCTCAGACCC 933  
Qy 56 GlnGlyGlyLeuTyrThrPheAlaProTyrSerValIleIleProGlnGly----- 72  
Db 934 CAAGAC-----CAGTTCAATATCATTTGTTCACTGCGGAGCAAC 975  
Qy 73 SerTPaenserCyValAlaGluCysAlaValAsnThrIleLysSerAspLeuGluIle 92  
Db 976 CAGTGGAGCAGCTGCTGTCGACACACAGAGAACTTGAACAGCGCGTTGACTAT 1035  
Qy 93 PheGlyAspLeuThrProValGlyAspGlyIleLysMetHisGlnThrValIle----- 110  
Db 1036 GCTTCCAAGATC---CCGGCTCAGGAGGAGCAACATCAATAAGCAGTCTATCGGCT 1092  
Qy 111 -----AsnGln-----MetProGlnAlaAla----- 118  
Db 1093 GTGGAACGCTGATTAAGCAACAGCGCTAGCTAGCTCCCAAGAGCGTTCCCTC 1152  
Qy 119 ValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsnProValGluGluValLys 138



```

Db 1153 ATCATCTGCTCAGGATGGCGAGCCCACTGTGGGG----- 1188
Qy 139 SerIleTyrGlnThrAspProAsnVal 147
Db 1189 -----GAGACCAATCCCAAGATT 1206

RESULT 18
AAK61746/c
ID AAK61746 standard; DNA; 1626 BP.
XX
AC AAK61746;
XX
DT 19-JUN-1999 (first entry)
XX
DE B. burgdorferi antigenic protein coding sequence, t867.nt.
XX
KM Antigenic protein; vaccine; Lyme disease; infection; detection; ss.
XX
OS Borrelia burgdorferi.
XX
PN MO9859071-A1.
XX
PD 30-DEC-1998.
XX
PF 18-JUN-1998; 98MO-US12718.
XX
PR 03-SEP-1997; 97US-0057483.
PR 20-JUN-1997; 97US-0050359.
PR 22-JUL-1997; 97US-0053344.
PR 22-JUL-1997; 97US-0053377.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (MEDI-) MEDIMUNE INC.
XX
PI Choi GH, Erwin AL, Hanson MS, Lathigra R;
XX
DR WPI; 1999-189980/16.
DR P-PSDB; AAY20049.
XX
PT New isolated Borrelia burgdorferi nucleic acids - used to develop
PT products for the diagnosis, prevention and treatment of diseases
PT caused by Borrelia, particularly Lyme disease
XX
PS Claim 1; Page 174; 275pp; English.
XX
CC This sequence encodes a Borrelia burgdorferi (Bb) protein of the
CC invention, which is suitable for use in a vaccine. The Bb polypeptides
CC can be used in vaccines for eliciting protective antibodies to members of
CC the Borrelia genus, particularly for the use against Lyme disease in
CC humans and animals. They can be used for preventing or attenuating an
CC infection caused by a member of the Borrelia genus. The products can also
CC be used for detection of members of the Borrelia genus.
XX
SQ Sequence 1626 BP; 508 A; 217 C; 355 G; 546 T; 0 other;

Alignment Scores:
Pred. No.: 1.64 Length: 1626
Score: 87.00 Matches: 45
Percent Similarity: 46.00% Conservative: 24
Best Local Similarity: 30.00% Mismatches: 49
Query Match: 7.77% Indels: 32
DB: 20 Gaps: 10

US-10-034-500-2 (1-218) x AAK61746 (1-1626)
Qy 20 TyrSerGlySerMetMetCysHisValAlaValArgGluProIleGluLeuAla 39
Db 986 TATGCCGGAAGACCTCTCCGCGAGATTTCTCCAGCGGTCCAGACATTCTTCATT 927
Qy 40 Lys-GluAlaIleLeuLysIleAsnAlaAlaMetProLysMetSerTyrGlnIleGlyLe 59
Db 926 GCTTGAGCCCATCTTGAAATTGAAT-----CTGCCAAAAGAAATATCAAGGC----- 878

```

```

Qy 59 uTyrThrPheAlaProTyrSer-ValIleIleProGlnGly----- 72
Db 877 -----CATTTGGCTGTAAATATCATCCCAATATGTAATAGCATATAACAGAA 831
Qy 73 -----SerTrpAsn---SerCysValAlaGlu-CysAlaValAsnThrIle 86
Db 830 GCTTCTAGCTGCAACGATGAGATGATATTAACAATATATCAAGTCTTGCATT 771
Qy 87 LysSerAspLeuGluIlePheGlyArgLeuThrProValGlyAspGlyIleLysMetHis 106
Db 770 AAA---GATTCCCGGATTGTCATTAATTCG---GGAAATCTTTAAGACTT--- 720
Qy 107 GluThrValIleAsnGlnMetProGlnAlaAlaValIleLeuLeuThrAspGly-Hi 126
Db 719 TCTACCACTTCTCTGCTGCTGCTACACAGCTGCAATATATC-----ACTAATCAACA 666
Qy 126 SasAsnLeuGlyMetAsnProValGluGluValIleLeuSerIleTyrGlnThrAspAs 146
Db 665 TCAGCATTTGCACTTGTAACCTGCTGAAGAACCGTTTTCCTGCTCAAAAGACCCGGA 606
Qy 146 nValCys-----PheHisVal 151
Db 605 A---TGCAAAAAGTTCCACCTT 587

RESULT 19
AAH00718/c
ID AAH00718 standard; DNA; 1656 BP.
XX
AC AAH00718;
XX
DT 24-JUL-2001 (first entry)
XX
DE Borrelia burgdorferi nucleotide sequence SEQ ID NO:709.
XX
KM Species specific; genus specific; family specific; probe; detection;
KM identification; algal; archael; bacterial; fungal; parasitica;
KM microorganism; diagnosis; translation elongation factor Tu; toxin;
KM translation elongation factor G; RecA recombinase; resistance;
KM catalytic subunit of proton-translocating ATPase; antimicrobial;
KM vaccine; primer; ds.
XX
OS Borrelia burgdorferi.
OS WO200123604-A2.
XX
PN 05-APR-2001.
XX
PF 28-SEP-2000; 2000MO-CA01150.
XX
PR 28-SEP-1999; 99CA-2283458.
PR 19-MAY-2000; 2000CA-2307010.
XX
PA (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
XX
PI Bergeron MG, Boissinot M, Hulstsky A, Menard C, Ouellette M;
PI Picard FJ, Roy PH;
XX
WPI; 2001-245006/25.
XX
PT Nucleic acid sequences are used to generate universal probes and
PT primers which can be used to identify and detect the presence of algal,
PT archael, bacterial, fungal and parasitcal species in a test sample -
XX
PS Disclosure; Page 806-807; 1580pp; English.
XX
CC The present invention describes a method for generating a repertory of
CC nucleic acids of tuf, fus, atpD and/or recA genes from which probes
CC and/or primers are derived. The method comprises amplifying the nucleic
CC acids of determined algal, archael, bacterial, fungal and parasitcal
CC species with a combination of defined primer pairs. The method can be
CC used for producing probes and/or primers for detecting one or more
CC related microorganisms e.g. algae, archaea, bacteria, fungi and
CC parasites, for universal detection and for specific and ubiquitous

```





CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human  
 CC diagnostic coding sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

SO Sequence 1404 BP; 288 A; 384 C; 420 G; 312 T; 0 other;

Alignment Scores:  
 Pred. No.: 1.53 Length: 1404  
 Score: 86.50 Matches: 35  
 Percent Similarity: 34.03% Conservative: 14  
 Best Local Similarity: 24.31% Mismatches: 62  
 Query Match: 7.73% Indels: 33  
 DB: 23 Gaps: 6

US-10-034-500-2 (1-218) x AAS92979 (1-1404)

QY 62 PheAlaProTyrSerValIleIleProGlnGlySerTyrPheSerCysValAla---Glu 60

Db 1314 TTTTACCGTTTCTGTTCTGTCGCGAAGAGATGAGACCTCTTCAGAGATTCA 1255

QY 81 CysAlaValAenThrIleLysSerApleGluIlePheGlyArgLeuThrProValGly 100

Db 1254 CCCGTATCGGCAATTGACGGCGATCGAATATCTGCGTGGAAACCGCG----- 1201

QY 101 AspGlyIleLysMetHisGluThrValIleAenGlnMetProProGlnAlaAlaValIle 120

Db 1200 -----CAACAGAAATCACTTCACACATTAACCCATCCACCGCGATTGAC 1153

QY 121 LeuLeuThrAerGlyHisAenAenLysGlyMetAenProValGluGluValLysSerIle 140

Db 1152 ACGACCAATGCTGAAAGGCGCATCTGAGCATCGGT----- 1117

QY 141 TyrGlnThrAenProAenValCysPhe-----HisValValSerPhe----- 154

Db 1116 -----GAACCTCTCTACTCTTTCCGCGATCGAGCAGATTCGCGATTCGCCGCG 1066

QY 155 AlaAerAerAerAerGluGlyLysAlaIleIleAerGlnIleValAlaLysAenSerGlySer 174

Db 1065 GCTCCTGACGCC-----CTTCGCCAATAATGTTTCAGCA 1033

QY 175 ValLeuValAerGlyLysGlnLeuLeuLysAenProAlaValCysGlnGluPheValAen 194

Db 1032 ATACAGACGACGGAAGTTTGGCTTCGCGAAGCGCCCATTCGATTCGATTCGCG 973

QY 195 SerValPheCys 198

Db 972 CTCATATCTCTGC 961

RESULT 23  
 ID AAA52180 standard; DNA; 3667 BP.

AC AAA52180;

XX 11-SEP-2000 (first entry)

DE DNA encoding H. influenzae strain K1 mature hmwA, SEQ ID NO:36.

KM Mature HMW protein; hmw gene; hmwA1; hmwA2; high molecular weight;

KW non-typeable Haemophilus influenzae; NTHI; non-encapsulated; recombinant production; Escherichia coli; antibacterial; vaccine;

KW human disease; otitis media; epiglottitis; pneumonia; tracheobronchitis;  
 KW detection; diagnosis; ds.

OS Haemophilus influenzae strain K1.

EH Key Location/Qualifiers

FT CDS 1..3667

FT /cag= a

FT /partial

FT /product= "Haemophilus influenzae strain K1 mature

FT /note= "No initiation codon given in the specification"

FT /transl\_except= (pos:919..920, aa:Gly)

FT 1..941

FT /\*cag= b

FT /partial

FT /product= "Haemophilus influenzae strain K1 mature

FT /note= "No initiation or termination codon given in the specification"

FT MO200020609-A2.

PD 13-APR-2000.

PF 07-OCT-1999; 99WO-CA00938.

PR 07-OCT-1998; 98US-0167568.

PR 08-DEC-1998; 98US-0206942.

PA (CONN-) CONNAUGHT LAB LTD.

PI Loosmore SM, Yang Y, Klein MH;

DR WPI; 2000-303789/26.

DR P-PSDB; AAB01830, AAB01831.

PT Nucleic acid molecule for producing recombinant high molecular weight

PS protection against Haemophilus induced diseases in humans -

PS Claim 7; Fig 20A-R; 307pp; English.

CC The invention relates to the recombinant production of Haemophilus  
 CC influenzae high molecular weight (HMW) proteins in Escherichia coli. The  
 CC expression construct used to effect recombinant expression comprises a  
 CC promoter functional in E. coli (e.g., the T7 promoter) operably linked  
 CC to a modified hmwABC operon from a non-tyeable (non-encapsulated) H.  
 CC influenzae (NTHI). Most HMW-expressing NTHI strains contain two hmw gene  
 CC clusters termed hmw1ABC and hmw2ABC. Each hmwABC operon comprises hmwA,  
 CC hmwB and hmwC genes. The hmwA genes encode the structural HMWA proteins  
 CC and the hmwB and hmwC genes encode accessory proteins which are  
 CC responsible for post-translational processing and secretion of the HMWA  
 CC proteins. The modified hmwABC operon used in the expression construct of  
 CC the invention contains an A gene modified such that it encodes only the  
 CC mature HMWA. The invention also discloses hmwA genes (AAA52175-A52198)  
 CC and HMWA proteins (AAB01824-B01849) from the non-tyeable H. influenzae  
 CC strains JcyC, K1, K21, LCWC2, PMH1, 15 and 12. The nucleic acids and  
 CC vectors are used for the production of recombinant H. influenzae HMW  
 CC proteins which can be used as vaccines to mediate a humoral or  
 CC cell-mediated immune response to provide protection against diseases in  
 CC humans caused by H. influenzae (e.g., otitis media, epiglottitis,  
 CC pneumonia and tracheobronchitis). The HMW proteins are also useful as  
 CC antigens in immunoassays for detecting antibodies against Haemophilus,  
 CC HMW proteins and/or HMW peptides. The nucleotide sequences encoding the  
 CC HMW proteins can be used to isolate and clone hmw genes from other  
 CC non-tyeable strains of Haemophilus via hybridisation reactions. The  
 CC present sequence represents a modified NTHI hmwA gene encoding a mature  
 CC HMWA protein.

SO Sequence 3667 BP; 1347 A; 736 C; 722 G; 862 T; 0 other;

Alignment Scores:

```

Pred. No.: 7.19 Length: 3667
Score: 86.00 Matches: 48
Percent Similarity: 31.85% Conservatave: 38
Best Local Similarity: 17.78% Mismatches: 98
Query Match: 7.69% Indels: 86
DB: 21 Gaps: 7

US-10-034-500-2 (1-218) x AAs52180 (1-3667)

OY 11 ValGluSerTyrAsnTyrLeuValAspTyrSerGlySerMetMetLeuValAla 30
   |||||
DB 983 ATAGAGTCAACAAAGCTTATGCTCCAGAGGGTCAAGGCTAAATTCAAAGCGAAGT 1042
   |||||
OY 31 ValArgGluProLeuIleGluLeuAlaValAlaIleLeu----- 44
   |||||
DB 1043 TCGACACACGCGGCTTTCATTAATAAATGATTTATTTAAATCCATGGGGCAT 1102
   |||||
OY 45 -----LysIleAsnAlaIleMet 50
   |||||
DB 1103 ATATCATTCAGCAAGTTCAGGTATGATGATATCTCAAAAAGCCCTTATAGCCAT 1162
   |||||
OY 51 ProLysMetSerTyrGlnGlnGlyLeuTyrThrPheAla----- 63
   |||||
DB 1163 AAAAACATTAACCTTGAAGGGGCAATATCACCCCTTGACGCGATAAACCAATAGA 1222
   |||||
OY 64 ---ProTyrSerValIleIleProGlnGlySerTyrPAsnSerCysValAlaGluCysAla 82
   |||||
DB 1223 ATCAAAAGTAAATATTTCTGTAAAGAGAGCC---AATGTCAACCTTGTAGCGGCAAT 1279
   |||||
OY 83 ValAsnThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGlyAspGly 102
   |||||
DB 1280 TATGTATATGACAAATACGCTTAAAGTATAGAGAAATGTCTACTAATAAGCAAT--- 1336
   |||||
OY 103 IleLysMetIleGluThrValIleAsnGlnMetProGlnAlaIleValIleLeuLeu 122
   |||||
DB 1337 -----CTCACCGT 1345
   |||||
OY 123 ThrAspGlyHisAsnAsnLeuGlyMetAsnProValGluGluValLysSerIleTyrGln 142
   |||||
DB 1346 ACCGGCTCCGCTATCAATATAGAAAAAATCTTACCGCTGAAAGTGTAGTATTTTA 1405
   |||||
OY 143 ThrAsnProAsnValCysPheHisValIleSerPheAlaAspAsp----- 157
   |||||
DB 1406 GCTAATCAATCAATATACAGCTTTAAGCTTCGCGCTTTCAGCAACCAAGCAAGTCAAC 1465
   |||||
OY 158 -----AlaGluGlyLysAlaIleIleAspGlnIleValAlaLeuAsnSerGlySer 174
   |||||
DB 1466 ATTTCCATCGCTAAGGAGAGAGCTATTTTAAAGATATCGAG-----AATACTGGCAGT 1519
   |||||
OY 175 ValIleVal----- 177
   |||||
DB 1520 CTGAATATTAACCACTAAATCCGATCCCAACCACTATTTATAAGGTAATTAAGT 1579
   |||||
OY 178 -----AspGlyLeuGlnLeuLeuGlnAsn 185
   |||||
DB 1580 AACAGAAAGAGTATTAAATATCACGAATATAGTATGATTAAGTCAAAATTTGGC 1639
   |||||
OY 186 ProAlaValCysGlnGluPheValAsnSerValPheCysGlnGluGlnIleLeuValThr 205
   |||||
DB 1640 GGCATATATCTCGCAAAAGAGCAATCTCACAATTTCTTGATTAAGTCAATATTAAC 1699
   |||||
OY 206 GluGluValValValLeuArgGlyValAsn 215
   |||||
DB 1700 GACCGGATTAACAATCAAGCAGCGCTTAAT 1729
   |||||

RESULT 24
ID AAA52179 strand; DNA; 3685 BP.
AC AAA52179;
XX
XX
XX 11-SEP-2000 (first entry)
XX

```

```

DE Haemophilus influenzae strain KI defective hmwA gene, SEQ ID NO:33.
XX
XX HMW protein; hmw gene; hmwA1; hmwA2; high molecular weight;
XX non-typable Haemophilus influenzae; NTHi; non-encapsulated;
XX recombinant production; Escherichia coli; antibacterial; vaccine;
XX human disease; otitis media; epiglottitis; pneumonia; tracheobronchitis;
XX detection; diagnosis; ds.
XX
XX Haemophilus influenzae strain KI.
OS
XX
XX
XX Key Location/Qualifiers
XX CDS 1..3685
XX /tag= a
XX /partial
XX /product= "Haemophilus influenzae strain KI full-length
FT /note= "No initiation codon given in the specification"
FT /transl_except= (pos:937..938, aa:Gly)
FT 1..959
FT CDS /tag= b
FT /partial
FT /product= "Haemophilus influenzae strain KI truncated
FT /note= "No initiation codon given in the specification"
FT /transl_except= "No initiation or termination codon given in the
FT specification"
XX
XX MO200020609-A2.
XX
XX 13-APR-2000.
XX
XX 07-OCT-1999; 99WO-CA00938.
XX
XX 07-OCT-1998; 98US-0167568.
XX
XX 08-DEC-1998; 98US-0206942.
XX
XX (CONN-) CONNAUGHT LAB LTD.
XX
XX Loesmore SM, Yang Y, Klein MH;
XX
XX WPI: 2000-303789/26.
XX
XX P-PSDB; AAB01828, AAB01829.
XX
XX Nucleic acid molecule for producing recombinant high molecular weight
XX proteins of Haemophilus which are used as a vaccine to provide
XX protection against Haemophilus induced diseases in humans -
XX
XX Claim 12; Fig 20A-R; 307pp; English.
XX
XX The invention relates to the recombinant production of Haemophilus
XX influenzae high molecular weight (HMW) proteins in Escherichia coli. The
XX expression construct used to effect recombinant expression comprises a
XX promoter functional in E. coli (e.g., the T7 promoter) operably linked
XX to a modified hmwABC operon from a non-typable (non-encapsulated) H.
XX influenzae (NTHi). Most HMW-expressing NTHi strains contain two hmw gene
XX clusters termed hmw1ABC and hmw2ABC. Each hmwABC operon comprises hmwA,
XX hmwB and hmwC genes. The hmwA genes encode the structural HMWA proteins
XX and the hmwB and hmwC genes encode accessory proteins which are
XX responsible for post-translational processing and secretion of the HMWA
XX proteins. The modified hmwABC operon used in the expression construct of
XX the invention contains an A gene modified such that it encodes only the
XX mature HMWA. The invention also discloses hmwA genes (AAs52175-A52198)
XX and HMWA proteins (AAB01824-B01849) from the non-typable H. influenzae
XX strains Joyce, KI, K21, LCDC2, PMH1, 15 and 12. The nucleic acids and
XX vectors are used for the production of recombinant H. influenzae HMW
XX proteins which can be used as vaccines to mediate a humoral or
XX cell-mediated immune response to provide protection against diseases in
XX humans caused by H. influenzae (e.g., otitis media, epiglottitis,
XX pneumonia and tracheobronchitis). The HMW proteins are also useful as
XX antigens in immunoassays for detecting antibodies against Haemophilus,
XX HMW proteins and/or HMW peptides. The nucleotide sequences encoding the
XX HMW proteins can be used to isolate and clone hmw genes from other
XX non-typable strains of Haemophilus via hybridisation reactions. The
XX present sequence represents an hmwA gene from a non-typable strain of H.

```

CC Influenzae.

XX Sequence 3685 BP; 1353 A; 737 C; 728 G; 867 T; 0 other;

Alignment Scores:

Pred. No.:	7.25	Length:	3685
Score:	86.00	Matches:	48
Percent Similarity:	31.85%	Conservative:	38
Best Local Similarity:	17.78%	Mismatches:	98
Query Match:	7.69%	Indels:	86
DB:	21	Gaps:	7

US-10-034-500-2 (1-218) x AAA52179 (1-3685)

```

QY 11 ValGuseTyrAsnTyrLeuValAspTyrSerGlySerMetMetLeuHisValAla 30
   |||||
Db 1001 ATAGAGTCACAAACTTAGTCCCTCAGAGGCTCAACCTTAAATTCAAAAGCGAAGT 1060
QY 31 ValArgGluProLysIleGluLeuAlaLysGluAlaIleLeu----- 44
Db 1061 TCGACACAGCCGCTTTACATATAAAATGATTTAATTTAAATGCGACTGGGGCAAT 1120
QY 45 -----LysIleAsnAlaIleMet 50
Db 1121 ATATCATGAACCAAGTTCAGAGTATGATGATGATCTCAAAAAGCCTTATAGCCAAT 1180
QY 51 ProLysMetSerTyrGlnGlyGlyLeuTyrThrPheAla----- 63
Db 1181 AAAAACAATACCTTGAAGGGGGCAATATCACCTTCGACCGCATAAAACCAATAGAA 1240
QY 64 --ProTyrSerValIleIleProGlnIySerTyrPheAsnSerCysValAlaGluCysAla 82
Db 1241 ATCAAAAGGTAATATATACGTTAAAGAGAGGCC--AATGTCACCTTCGTAGCGCAAT 1297
QY 83 ValAsnThrIleLysSerAspLeuGluIlePheGlyAlaGlyLeuThrProValGlyAspGly 102
Db 1298 TATGGTAATGACAAATTCAGCTTTAAGTAAAGGAAATGTCATCAATAAAGGCAAT-- 1354
QY 103 IleLysMetHisGluThrValIleAsnGlnMetProGlnAlaAlaValIleLeuLeu 122
Db 1355 -----CTCACCGTT 1363
QY 123 ThrAspGlyHisAsnAsnLeuGlyMetAsnProValGluGluValLysSerIleTyrGln 142
Db 1364 ACCGGCTCCGCTATCATATATAGAAAATAATCTTACCGTTGAAGGTAGCTAAAGTTTAA 1423
QY 143 ThrAsnProAsnValCysPheHisValValSerPheAlaAspAsp----- 157
Db 1424 GCTAATCAAAATTACAGCTTAAAGTATCCGGCTATTTGACAAACCAAGCAAGTCAAC 1483
QY 158 -----AlaGluGlyLysAlaIleIleAspGlnIleValAlaLeuAsnSerGlySer 174
Db 1484 ATTTCCATCGCTTAAGGAGAGAGCTATTTTAAAGATATCGAG-----AATATCGCAGT 1537
QY 175 ValLeuVal----- 177
Db 1538 CTGATATATTACACTAAATCCGACTCAACACCATATATTAAAGGTATATAAT 1597
QY 178 -----AspGlyLeuGlnLeuGlnAsn 185
Db 1598 AACAGAAAAGGTATTAAATATACGAAATATGATGATATATACCAATTCGATGGC 1657
QY 186 ProAlaValCysGlnGluPheValAsnSerValPheCysGlnGluGlnIleLeuValThr 205
Db 1658 GGCATATCTCGCAAAAGAAAGGCAATCTCACAAATTTCTTGTGATTAAGTCAATATTAC 1717
QY 206 GluGluValValLeuArgGlyValAsn 215
Db 1718 GAGCGATATACATCAAGACAGCGCTTAAT 1747

```

RESULT 25

AAI40781  
ID AAI40781 standard; DNA; 88421 BP.

XX	AAI40781;	XX	Location/Qualifiers
AC		XX	2077..3078
XX		XX	/*tag= a
DT	03-Oct-2002 (first entry)	XX	/product= "Protein of ORF 1"
XX		XX	3118..4032
XX	8842Int genomic DNA of ramoplanin producing Actinoplanes sp.	XX	/*tag= b
DE		XX	/product= "Protein of ORF 2"
XX		XX	4038..5048
XX		XX	/*tag= c
XX		XX	/product= "Protein of ORF 3"
XX		XX	complement (6665..5814)
XX		XX	/*tag= d
XX		XX	/product= "Protein of ORF 4"
XX		XX	complement (7703..6693)
XX		XX	/*tag= e
XX		XX	/product= "Protein of ORF 5"
XX		XX	complement (9464..8130)
XX		XX	/*tag= f
XX		XX	/product= "Protein of ORF 6"
XX		XX	9691..10761
XX		XX	/*tag= g
XX		XX	/product= "Protein of ORF 7"
XX		XX	complement (12751..10829)
XX		XX	/*tag= h
XX		XX	/product= "Protein of ORF 8"
XX		XX	complement (13617..12802)
XX		XX	/*tag= i
XX		XX	/product= "Protein of ORF 9"
XX		XX	complement (15203..13614)
XX		XX	/*tag= j
XX		XX	/product= "Protein of ORF 10"
XX		XX	15591..15863
XX		XX	/*tag= k
XX		XX	/product= "Protein of ORF 11"
XX		XX	15880..19035
XX		XX	/*tag= l
XX		XX	/product= "Protein of ORF 12"
XX		XX	19032..39713
XX		XX	/*tag= m
XX		XX	/product= "Protein of ORF 13"
XX		XX	39713..65800
XX		XX	/*tag= n
XX		XX	/product= "Protein of ORF 14"
XX		XX	65826..66530
XX		XX	/*tag= o
XX		XX	/product= "Protein of ORF 15"
XX		XX	66546..67370
XX		XX	/*tag= p
XX		XX	/product= "Protein of ORF 16"
XX		XX	67384..70059
XX		XX	/*tag= q
XX		XX	/product= "Protein of ORF 17"
XX		XX	70099..70662
XX		XX	/*tag= r
XX		XX	/product= "Protein of ORF 18"
XX		XX	70659..71906
XX		XX	/*tag= s
XX		XX	/product= "Protein of ORF 19"
XX		XX	complement (73439..71964)

FT		/tag= t
FT		/product= "Protein of ORF 20"
FT	CDS	/complement (74216..73563)
FT		/tag= u
FT		/product= "Protein of ORF 21"
FT	CDS	/complement (75424..74213)
FT		/tag= v
FT		/product= "Protein of ORF 22"
FT	CDS	/complement (7535..76464
FT		/tag= w
FT		/product= "Protein of ORF 23"
FT	CDS	/complement (78110..76449)
FT		/tag= x
FT		/product= "Protein of ORF 24"
FT	CDS	/complement (79864..78107)
FT		/tag= y
FT		/product= "Protein of ORF 25"
FT	CDS	/complement (81624..79861)
FT		/tag= z
FT		/product= "Protein of ORF 26"
FT	CDS	/complement (81909..81682)
FT		/tag= aa
FT		/product= "Protein of ORF 27"
FT	CDS	/complement (82346..82062)
FT		/tag= ab
FT		/product= "Protein of ORF 28"
FT	CDS	/complement (82587..84446
FT		/tag= ac
FT		/product= "Protein of ORF 29"
FT	CDS	/complement (84481..85548
FT		/tag= ad
FT		/product= "Protein of ORF 30"
FT	CDS	/complement (85556..86845
FT		/tag= ae
FT		/product= "Protein of ORF 31"
FT	CDS	/complement (87372..86803
FT		/tag= af
FT		/product= "Protein of ORF 32"
FT	CDS	/complement (87494..88420
FT		/tag= ag
FT		/product= "Protein of ORF 33"
FN		
PN		WO200231155-A2.
PD		
PD	18-APR-2002.	
PX		
PX		
PF	15-OCT-2001; 2001WO-CA01462.	
PR	13-OCT-2000; 2000US-239924P.	
PR	12-APR-2001; 2001US-283296P.	
PR	24-JUL-2001; 2001US-0910813.	
XX		
PA	(ECOP-) ECOPIA BIOSCIENCES INC.	
XX		
PI	Farnet CM, Zakopoulos E, Staffa A;	
XX		
XX	WPI: 2002-435445/46.	
DR	P-PSTDB: AAO22146, AAO22147, AAO22148, AAO22149, AAO22150, AAO22151,	
DR	AAO22152, AAO22153, AAO22154, AAO22155, AAO22156, AAO22157, AAO22158,	
DR	AAO22159, AAO22160, AAO22161, AAO22162, AAO22163, AAO22164, AAO22165,	
DR	AAO22166, AAO22167, AAO22168, AAO22169, AAO22170, AAO22171, AAO22172,	
XX	AAO22173, AAO22174, AAO22175, AAO22176, AAO22177, AAO22178.	
PT		
PT	Novel isolated ramoplanin biosynthetic pathway polypeptide useful for	
PT	chemically modifying biological molecule that is a substrate for a	
PT	polypeptide encoded by a ramoplanin biosynthesis gene cluster -	
PS		
XX	Disclosure; Page 87-135; 212pp; English.	
CC	The invention relates to an isolated ramoplanin biosynthetic pathway	
CC	polypeptide selected from a polypeptide of open reading frames (ORF)	
CC	1-32. The isolated polypeptides are useful for chemically modifying a	
CC	biological molecule that is a substrate for a polypeptide encoded by a	

[illegible]



```

RESULT 26
AB069245/C
ID AB069245 standard; DNA; 3011208 BP.
XX
AC AB069245;
XX
DT 29-AUG-2002 (first entry)
XX
DE Listeria innocua DNA sequence #684.
XX
KM Antibacterial; Listeria; food contamination; mutational analysis;
XX infection; ds.
XX
OS Listeria innocua.
XX
PN WO200228891-A2.
XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-FR03061.
XX
PR 04-OCT-2000; 2000FR-0012697.
XX
PA (INSP ) INST PASTEUR.
XX (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI Kunst F, Glaser P;
XX
DR WPI; 2002-332479/37.
XX
XX New genomic sequences from Listeria species, useful for detection,
PT treatment and prevention of infection, also related polypeptides,
XX antibodies and modulators -
XX
XX Claim 5; SEQ ID 2058; 180bp; French.
XX
XX The present invention relates to nucleic acid sequences
XX (AB067188-AB071212) from Listeria sp. The sequences are useful as probes
XX and primers for identification and/or detection of Listeria (e.g. as
XX contaminants in foods, or mutational analysis) and for analysis of
XX gene expression. Proteins encoded by the nucleic acid sequences can be
XX used to screen for compounds that modulate gene expression, replication
XX and pathogenicity of Listeria (potential therapeutic agents), also for
XX treating infections by Listeria, and are useful as immunogens in
XX anti-Listeria vaccines.
XX Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 3011208 BP; 941651 A; 568176 C; 559189 G; 942192 T; 0 other;
XX
Alignment Scores:
Pred. No.: 1.31e+05 Length: 3011208
Score: 86.00 Matches: 55
Percent Similarity: 35.69% Conservative: 36
Best Local Similarity: 21.57% Mismatches: 76
Query Match: 7.69% Indels: 88
DB: 24 Gaps: 14
US-10-034-500-2 (1-218) x AB069245 (1-3011208)
QY 21 SerGlySerMetMetMetLeuValAlaValArgGluProLysIleGluLeuAlaLys 40
DB 2563623 AATGGGAATTTATGACGCGCAATGGCTATGATTTACAGAA----- 2563579
QY 41 GtualaileuLysIleAsnAlaIaIaMetProLysMetSerTyrGln---GlyGlyLeu 59
DB 2563578 -----GATATTAAATACGCGTAATTAGATGACAATATATGTTGGGAGAT 2563531
QY 60 TyrThrPheAlaProTyr-----SerValIleIleProGln 71
DB 2563530 TTAAGTGTAGCAGCAATTGTACAAAGAGCCAAACAGAAAAAGATTATTATACAGAA 2563471

```

```

QY 72 GlySerTTPAsnSerCysValAlaGluCysAlaValAsnThrIleLysSerAspLeuGlu 91
DB 2563470 GCGCAATGC-----ATGAT 2563456
QY 92 IlePhe---GlyValGluThrProValGlyAspGlyIleLysMetHisGluThrValIle 110
DB 2563455 ATTTGGAACGGCGGATTCATCCAGTGGCGAAACAAATTTCTTACTATGCAGATGTT--- 2563399
QY 111 AsnGlnMetPro-----ProGlnAlaIaValIleLeu-----LeuThrAspGly 125
DB 2563398 GATACGCTACCTGTATTGCAAAAGCTGAGCCGATTATTCATGAATGAATGACAGATGCT 2563339
QY 126 HisAsnAsnLeuGlyMetAsnProValGluValIleLysSerIleTyrGlnThrAsnPro 145
DB 2563338 TAC---CACTAGCTCAAAATGTATGAGAAAGCATTTAAAGCATAT-----GAT 2563294
QY 146 AsnValCysPheHisVal-----ValSerPheAlaAspAspAla 158
DB 2563293 AATTTAAGTTCCGAGTTTACCCATCAGAAACAGCAGATGACGTTTATGATGATGTA 2563234
QY 159 GluGlyLysAlaIleIleAsp----- 165
DB 2563233 AATGTGGCGAGATGCGTATATTAGCGTTTGAAGACTTTGCAAAATGAAAAATTACG 2563174
QY 166 -----GlnIleValAlaLeuAsnSerGly 173
DB 2563173 GTCAATCTCCGCGATGCGAGATGAAACAAACAGATGCAAGTTTCAATACAGACCACA 2563114
QY 174 SerValIleuValAspGlyLeuGlnLeuLeuGln-----AsnPro 186
DB 2563113 AGTGTACAGATGCGAGAGCTAGCTAGTACAGACACACTAGATGCTTTAAATGAA 2563054
QY 187 AlaValCysGlnGluPheValAsnSerVal-----PheCysGlnGluGlnIle 202
DB 2563053 GCACAAACCGCTTATTTACTAGACACTGACCAAACTTAACCTTACATCAAAATGCGGCA 2562994
QY 203 LeuValThrGluGluValValIleValArgGlyValAsnPheAla 217
DB 2562993 ACAGATCTACTACAGCTATCGTTTAAATGCTAAATCATGCG 2562949
RESULT 27
AAZ51000
ID AAZ51000 standard; DNA; 3720 BP.
XX
AC AAZ51000;
XX
XX 05-JUN-2000 (first entry)
XX
DE X. oryzae avrXol gene.
XX
KM avrXol gene; Xanthomonas oryzae pv. oryzaicola; avirulence product;
XX hypersensitive resistance; transgenic plant; cell death; abscission;
XX fruit harvesting; herbicide development; se.
XX
OS Xanthomonas oryzae.
XX
FH Key Location/Qualifiers
FT CDS 1..3720
FT /tag= a
FT /product= "avrXo protein"
XX
PN WO200009698-A2.
XX
PD 24-FEB-2000.
XX
PF 14-AUG-1999; 99WO-US18588.
XX
PR 14-AUG-1998; 98US-0096585.
XX
PR 13-AUG-1999; 99US-0374273.
XX
XX (UNIV ) UNIV KANSAS STATE RES FOUND.
XX

```



PI Leach JE, Hubert SH, Ardales E;  
 XX WPI: 2000-22437/19.  
 DR P-PSDB; AAY70062.  
 XX Novel avirulence gene from *Xanthomonas oryzae*, used to generate  
 PT transgenic plants that show e.g. inducible resistance properties or  
 PT cell death  
 XX  
 XX Claim 1; Page 36-37, 46pp; English.  
 CC The present sequence is avrKx1 gene from the rice pathogen,  
 CC *Xanthomonas oryzae* pv. *oryzicola*. This gene encodes an avirulence  
 CC product that can cause a hypersensitive resistance response in maize.  
 CC Vectors containing avrKx1 gene, preferably under the control of a  
 CC tightly regulated inducible promoter, are used to produce transgenic  
 CC plants that have inducible resistance properties. The hypersensitivity  
 CC response in plants like maize, that recognise the gene product, includes  
 CC cell death, e.g. to prevent flowering, to induce abscission (to  
 CC facilitate fruit harvesting), for herbicide development and for killing  
 CC above-ground parts of plants before harvest.  
 XX  
 SQ Sequence 3720 BP, 646 A, 1192 C, 1295 G, 587 T, 0 other;  
 Alignment Scores:  
 Pred. No.: 8.48 Length: 3720  
 Score: 85.50 Matches: 58  
 Percent Similarity: 32.46% Conservative: 29  
 Best Local Similarity: 21.64% Mismatches: 84  
 Query Match: 7.64% Indels: 97  
 DB: 21 Gaps: 10  
 US-10-034-500-2 (1-218) x AA251000 (1-3720)  
 QY 31 ValArgGluProLysIleGluLeuAlaLysGluAlaIleLeuLysIleAsn----- 47  
 DB 742 TTGAGAGGTCCTCGCTTACAGTTGACACAGGCCAACCTTCTCAAGATTGCAGAAACGTGGC 801  
 QY 48 -----AlaAlaMetProLysMetSerTyrGlnGlyGlyLeuTyrThrPheAlaProTyr 65  
 DB 802 GCGGTGACCGGACGTGAGGACGATGCATGCATCGGACATGACATGACGGGTGCCCCCTG 861  
 QY 66 SerValIleIleProGlnGlySerTyrPasnSerCysValAlaGluCysAlaValAlaThr 85  
 DB 862 AACCTGACCCCGGCACAG-----GTGCTGCCATTCGACACATTAACGCG 906  
 QY 86 IleLysSerAspLeuGluIlePheGlyArgLeuThrProVal----- 99  
 DB 907 GGCACGACGCGCTGGAGACGTGACGCGGTGCTGCTGCTGCCAGACATGCG 966  
 QY 100 -----GlyAspGlyIleLysMetHisGlu 107  
 DB 967 CTGACCCCGGACGAGCTGTCGTCATGCCAGCATGCGCGGCGGACAGCGGCTGGAG 1026  
 QY 108 ThrValIleAsnGlnMet-----ProProGlnAla 117  
 DB 1027 ACGGTGACGCGGTGTCGCGGTGCTGTGCCAGGACCATGACCTGACCCGACACAGTGG 1086  
 QY 118 AlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsnProValGluGluVal 137  
 DB 1087 GTGGCCATCCCGACGACGATGCG-----GGCAAGCAGCGCGTGGAGACGCTG 1134  
 QY 138 LysSerIleTyrGlnThrAsnProAsnValCys-----PheHis 150  
 DB 1135 CACGCGGTGTG-----CCGGTCTGTGTCGACGACCATGCGCTGACCCCGGACAG 1185  
 QY 151 ValValSerPheAlaAspAspAlaGluGlyValAlaIleIle----- 164  
 DB 1186 GTGTGTCATTCGCGACGATGGCGGGGCAAGCAGCGCTGGAGAGCGGTGACGCGCTG 1245  
 QY 165 -----AspGlnIleValAlaLeuAsnSer 172  
 DB 1246 TTGCGGTGCTGTGTCAGACCATGCTGACCCCGGACGAGGTGTGTCATGCGCAGC 1305

QY 173 GlySerValIleuValAspGlyLeuGlnLeuLeuGlnAsn-----ProAlaValCysGln 190  
 DB 1306 CACGATGGCGGACGACGAGCGCGCTGTGAGACGCGGTGTCGCGGTCTGTGCGAG 1365  
 QY 191 GluPhe----- 192  
 DB 1366 GACCATGCGCTGACCCCGGATCAGGTGGCCATCGCCACGACGATGCGGACAGCAG 1425  
 QY 193 -----ValAsnSerValPheCysGlnGluGlnIleLeuValThr 205  
 DB 1426 GCGCTGAGAGCGGTGACGCGGTGTCGCGGTGCTGTGTCGACGACCATGCGCTGACCCG 1485  
 QY 206 GluGluValValValIleuArgGly 213  
 DB 1486 GACGACGTGCTGTGCGCGGTGCGCGGC 1509  
 RESULT 28  
 AA160237  
 ID AA160237 standard; cDNA; 6292 BP.  
 AC AA160237;  
 XX  
 DT 22-OCT-2001 (first entry)  
 XX  
 DE Human polynucleotide SEQ ID NO 4226.  
 XX  
 KW Human, noctropic, immunosuppressant, cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200153312-A1.  
 XX  
 PD 26-JUL-2001.  
 XX  
 PF 26-DEC-2000; 2000WO-US34263.  
 XX  
 PR 21-JAN-2000; 2000US-0488725.  
 XX  
 PR 25-APR-2000; 2000US-0552317.  
 PR 09-JUL-2000; 2000US-0598042.  
 PR 19-JUL-2000; 2000US-0620312.  
 PR 03-AUG-2000; 2000US-0653450.  
 PR 14-SEP-2000; 2000US-0662191.  
 PR 19-OCT-2000; 2000US-0693036.  
 PR 29-NOV-2000; 2000US-0727344.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AU, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
 XX  
 DR WPI: 2001-442253/47.  
 DR P-PSDB; AAM41081.  
 PT Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -  
 XX  
 PS Claim 1; SEQ ID NO 4226; 10078pp; English.  
 XX  
 CC The invention relates to human nucleic acids (AA157796-AA161369) and  
 CC the encoded polypeptides (AAM38642-AA442213) with noctropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilization of the activities such as: Immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, hemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukemias and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.

XX Sequence 6292 BP; 1818 A; 1391 C; 1595 G; 1487 T; 1 other;

# Alignment Scores:

Pred. No.:	Length:	Score:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
18.3	6292	85.50	44	35	85	45	8
Percent Similarity:		37.98%					
Best Local Similarity:		21.15%					
Query Match:		7.64%					

US-10-034-500-2 (1-218) x AAD38694 (1-6292)

Qy 15 AenTYrLeuValAepTYrSerGlySerMetMetLysHisValAlaValArgGlu--- 33

Db 3901 AGCTTCTGATTGGTCTCTGGTCAAGATGTTCTCAACATGTCGCTCAACAGAAAAA 3960

Qy 34 -----ProLysIleGluLeuAlaLysGluAlaIleLeu 44

Db 3961 TGATCCGAAGACGACGATGACCCGCTTTCAGCGATGAGGAGACCACTCCCTCC- 4019

Qy 45 LysIleAenAlaAlaMetPro-----LysMetSerTYrGlnGlyGlyLeuTYrThr 61

Db 4020 CTGATGACACGATCCAAACCCGCGCCAGTGAAGCCTGTATTACGGTGG- 4067

Qy 62 PheAlaProTYrSerValIleIleProGlnGlySerTPanSerCys-----Val 78

Db 4068 -----CAATATGGCCAGTGGTCTCCATGCCAAGTCAGAGAG 4103

Qy 79 AlaGluCysAlaValAenThrIleLysSerAepLeuGluIlePheGlyArgLeuThrPro 98

Db 4104 GCCACGTGTGGAGAGGAGACCGAAGCAAGAACATTCTCTGTGTA----- 4148

Qy 99 ValGlyAepGlyIleLysMetHisGluThrValIleAenGlnMetProProGlnAlaAla 118

Db 4149 GTAAGTATGGGTGACGATGATTTTCAGCAAGGTGGTGAAGAAATTCGTGTCGAC 4208

Qy 119 ValIleLeuLeuThrAenGlyHisAenLeuGlyMetAenProValGluGluValLys 138

Db 4209 ATGAACCTCATTTATGATGATGATATAATAATATGTTCTGGAG-----GAA 4253

Qy 139 SerIleTYrGlnThrAenProAenValCysPheHisValIleSerPheAlaAenAepAla 158

Db 4254 TCTCGACACCGACCTTCCGACGATGCTGT-----TATTTGAAGACCTGG 4298

Qy 159 GluGlyValAlaIleIleAepGlnIleValAlaLeuAenSerGlySerValLeuValAep 178

Db 4299 TCTTCTCGACACCTGTGTG---CAGCTGACCTGTGTGAATGCGACGATGAGCTTGTGT 4355

Qy 179 GlyLeuGlnLeuLeuGlnAenProAlaValCysGlnGlnIlePheValAenSerValPheCys 198

Db 4356 GGAATACAGGTCAGATCCAGACCGGTGATTTATCAAGAACTAGAGAAATCAGCATCTGTGC 4415

Qy 199 GlnGlnGlnIleLeuValThrGlu 206

Db 4416 CCAGGCGATGTTTGAACAACAAA 4439

RESULT 29

AAAD38694

XX AAD38694 standard; cDNA; 6317 BP.

XX AAD38694;

XX 23-SEP-2002 (first entry)

DE Human LP217 secreted protein encoding cDNA.

XX Human; secreted protein; atherosclerosis; Alzheimer's disease; LP217;  
 XX diabetic retinopathy; severe combined immunodeficiency; pancreatitis;  
 XX rheumatoid arthritis; colorectal adenoma; haemolytic anaemia; cancer;  
 XX reperfusion injury; arteriosclerosis; wound healing; transgenic animal;  
 XX gene therapy; neoplasm; transgenic; psoriasis; ischaemia; carcinoma;  
 XX chromosome 7p21-p22; gene; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 284..5158

XX /\*tag= a

XX /product= "Human LP217 secreted protein"

XX sig\_peptide 284..328

XX /\*tag= b

XX mat\_peptide 329..5155

XX /\*tag= c

XX /product= "Mature human LP217 secreted protein"

XX WO200226801-A2.

XX 04-APR-2002.

XX 14-SEP-2001; 2001WO-US26026.

XX 28-SEP-2000; 2000US-236088P.

XX (ELIL ) LILLY & CO ELI.

XX Su EW, Wang H;

XX WPI; 2002-471259/50.

XX P-PSDB; AAE23979.

XX Novel proteins and polynucleotides of secreted proteins useful for

XX treating various diseases e.g. rheumatoid arthritis, cancer, psoriasis,

XX diabetic retinopathy, arteriosclerosis, ischemia or reperfusion injury

XX Claim 1; Page 109-117, 145pp; English.

XX The invention relates to human secreted polypeptides designated LP095,

XX LP191, LP217, LP220, LP221, LP222, LP229, LP237 or LP238 and nucleic

XX acid molecules encoding such polypeptides. Novel secreted proteins of

XX the invention are used for treating diseases such as atherosclerosis,

XX Alzheimer's disease, diabetic retinopathy, psoriasis, pancreatitis,

XX arteriosclerosis, rheumatoid arthritis, colorectal adenoma, severe

XX combined immunodeficiency, ischemia, carcinomas, haemolytic anaemia,

XX reperfusion injury, neoplasms and cancer especially liver cancer. They

XX are also used for wound healing. Polynucleotides of the invention can

XX be used to generate transgenic animals or knock out animals, which in

XX turn, are useful in the development and screening of therapeutically

XX useful reagents for use in the treatment of diseases associated with

XX LP polypeptides associated activity. They are also used in gene therapy.

XX The present sequence is human LP217 secreted protein encoding cDNA.

XX LP217 gene is located on chromosome 7p21-p22.

XX Sequence 6317 BP; 1776 A; 1442 C; 1617 G; 1482 T; 0 other;

XX Alignment Scores:

XX Pred. No.:

XX 18.4

XX Length:

XX 6317

XX Score:

XX 85.50

XX Matches:

XX 44

XX Percent Similarity:

XX 37.98%

XX Conservative:

XX 35

XX Best Local Similarity:

XX 21.15%

XX Mismatches:

XX 85

XX Query Match:

XX 7.64%

XX Indels:

XX 45

XX Gaps:

XX 8

XX US-10-034-500-2 (1-218) x AAD38694 (1-6317)

Qy 15 AenTYrLeuValAepTYrSerGlySerMetMetLysHisValAlaValArgGlu--- 33

```

Db 4047 AGCTTTGATTGTCCTCCTGGTCAGATGTTCTCAAAACATGTGCTCACAGAAAAA 4106
Oy 34 -----ProlysiIleGluLeuAlaLysGluAlaIleLeu 44
Db 4107 TGATCCGAAGACGAACAGTGAACCCCTTTCAAGTGATGGAAGACCATCCCTTCC- 4165
Oy 45 LysIleAsnAlaIleMetPro-----LysMetSerTyrGlnGlyGlyLeuTyrThr 61
Db 4166 CTGATGACCAAGCTGCAAAACCTCCAGTGAAGCCTGTGTTATCGGTG- 4213
Oy 62 PheAlaProTyrSerValIleIleProGlnGlySerTyrPasnSerCys-----Val 78
Db 4214 -----CAATATGGCCAGTGGTCTCCATGCAAGTCAGAGGAG 4249
Oy 79 AlaGluCysAlaValAsnThrIleIleLysSerAspLeuGluIlePheGlyArgLeuThrPro 98
Db 4250 GCCCAGTGTGGAGAAAGGACCGAACAAGAACATTTCTTGTA----- 4294
Oy 99 ValGlyAspGlyIleLysMetHisGluThrValIleAsnGlnMetProProGlnAlaIle 118
Db 4295 GAAAGTGAAGTGTGACGCTGATGATTTACCAAAAGTGTGATGAGGAATCTGTGCTGAC 4354
Oy 119 ValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsnProValGluGluValLys 138
Db 4355 ATTGAACCTATTATAGATGTATATAAAATATGTTCTGGAG-----GAA 4399
Oy 139 SerIleTyrGlnThrAsnProAsnValCysPheHisValIleSerPheAlaAspAla 158
Db 4400 TCCTGCAGCCACCTTGGCCAGGTGACTGT-----TATTTGAAGGACTGG 4444
Oy 159 GluGlyValAlaIleLeuAspGlnIleValAlaLeuAsnSerGlySerValLeuValAsp 178
Db 4445 TCTTCTGAGACCTGTGT---CAGCTGACCTGTGTAAAGTGAGAGATCTAGGCTTGGT 4501
Oy 179 GlyLeuGlnLeuLeuGlnAsnProAlaValCysGlnGluPheValAsnSerValPheCys 198
Db 4502 GGAATACAGTGTGATGATCCAGACCGGTGATATATCAAGAACTAGAGATCAGCATGTGTC 4561
Oy 199 GlnGluGlnIleLeuValThrGlu 206
Db 4562 CCAGAGCAGATGTTAGAAACAAA 4585
RESULT 30
AAF30188
ID AAF30188 standard; cDNA; 6373 BP.
XX
AC AAF30188;
XX
DT 30-APR-2001 (first entry)
XX
DE Clone 21433858 encoding secreted protein SECP1.
XX
KW SECP1; secreted protein; human; diagnosis; gene therapy;
XX
KW lung cancer; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX FH 235..5001
XX FT /*tag= a
XX FT sig_peptide 235..303
XX FT /*tag= b
XX FT mat_peptide 304..4998
XX FT /*tag= c
XX
XX WO200105971-A2.
XX
XX PD 25-JAN-2001.
XX
XX PF 20-JUL-2000; 2000MO-US19890.
XX
XX PR 20-JUL-1999; 99US-0144722.
XX
XX PR 29-NOV-1999; 99US-0167785.

```

```

PR 19-JUL-2000; 2000US-0619252.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Shinkets RA, Fernandez E;
XX
XX WPI; 2001-091973/10.
XX
XX P-PSDB; AAB20155.
XX
XX New polypeptide designated SECP, its encoding nucleic acid and its
XX immunospecific antibody, useful for diagnosing, preventing and treating
XX SECP-associated disorders such as cancer.
XX
XX Claim 8; Fig 1; 124pp; English.
XX
XX The present sequence is that of clone 2143385 encoding novel
XX secreted protein SECP1 (see AAB20155) that is predicted to localise
XX in the plasma membrane. RNA species with homology to the clone
XX were detected in endometrial cells, pancreas, adipose, adrenal
XX gland, thyroid, mammary gland, myometrium, uterus, placenta,
XX prostate, testis, and in neoplastic cells derived from ovarian
XX carcinoma, breast carcinoma, prostate carcinoma (bone metastases)
XX and melanoma. SECP1 nucleic acids can be used to identify these
XX cell types. SECP1 was also highly expressed in a small cell lung
XX cancer, a large cell lung cancer, and a non-small cell lung
XX cancer. SECP1 nucleic acids may be useful as diagnostic probes for
XX such lung cancers, and nucleic acids and proteins may be targets
XX for the treatment of such cancers. The invention provides a novel
XX SECP secreted proteins (see AAB20155-63), nucleic acids encoding
XX them (see AAF30188-96), antibodies, mutants or fragments. These
XX can be used to detect, treat or prevent an SECP-associated disorder,
XX to screen for predisposition to such a disorder, and to identify an
XX agent that modulates the expression or activity of SECP.
XX
XX SQ Sequence 6373 BP; 1836 A; 1385 C; 1603 G; 1548 T; 1 other;
XX
XX Alignment Scores:
XX Pred. No.: 18.6 Length: 6373
XX Score: 85.50 Matches: 44
XX Percent Similarity: 37.98% Conservative: 35
XX Best Local Similarity: 21.15% Mismatches: 85
XX Query Match: 7.64% Indels: 45
XX DB: Gaps: 8
XX
XX US-10-034-500-2 (1-218) x AAF30188 (1-6373)
Oy 15 AsnTyrLeuValAspTyrSerGlySerMetMetLysHisValAlaValArgGlu--- 33
Db 3890 AGCTTTCTGATTGTCCTCCTGTGTCAGAAATGTTCTCAAAACATGTGCTCACAGAAAAA 3949
Oy 34 -----ProlysiIleGluLeuAlaLysGluAlaIleLeu 44
Db 3950 TGATCCGAAGACGAACAGTGAACCCCTTTCAAGTGATGGAAGACCATCCCTTCC- 4008
Oy 45 LysIleAsnAlaIleMetPro-----LysMetSerTyrGlnGlyGlyLeuTyrThr 61
Db 4009 CTGATGACCAAGCTGCAAAACCTCCAGTGAAGCCTGTGTTATCGGTG- 4056
Oy 62 PheAlaProTyrSerValIleIleProGlnGlySerTyrPasnSerCys-----Val 78
Db 4057 -----CAATATGGCCAGTGGTCTCCATGCAAGTCAGAGGAG 4092
Oy 79 AlaGluCysAlaValAsnThrIleIleLysSerAspLeuGluIlePheGlyArgLeuThrPro 98
Db 4093 GCCCAGTGTGGAGAAAGGACCGAACAAGAACATTTCTTGTA----- 4137
Oy 99 ValGlyAspGlyIleLysMetHisGluThrValIleAsnGlnMetProProGlnAlaIle 118
Db 4138 GAAAGTGAAGTGTGACGCTGATGATTTACCAAAAGTGTGATGAGGAATCTGTGCTGAC 4197
Oy 119 ValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsnProValGluGluValLys 138

```







CC cellular adhesion. The proteins can be used to treat any von Willebrand  
 CC factor-associated disorder, regulate extracellular matrix structuring,  
 CC cellular adhesion, and cell trafficking and/or migration, modulate  
 CC cellular interactions, modulate cell adhesion in proliferative  
 CC disorders, such as cancer, modulate the proliferation, differentiation,  
 CC and/or function of cells that appear in the bone marrow, and leukocytes,  
 CC treat bone marrow, blood and hematopoietic associated diseases and  
 CC disorders, atelectasis, pulmonary congestion or edema, emphysema,  
 CC chronic bronchitis, bronchial asthma and bronchiectasis, intestinal  
 CC disorders, spleen associated diseases, modulate renal disorders, treat  
 CC cardiovascular disorders such as ischemic heart disease, modulate the  
 CC proliferation, differentiation, and/or function of bone and cartilage  
 CC cells and to treat bone and/or cartilage associated diseases or  
 CC disorder. They may also be used to treat disorders associated with the  
 CC ovaries, cerebral edema, hydrocephalus, brain herniations, iatrogenic  
 CC diseases, inflammations, bacterial and viral meningitis, Alzheimer's  
 CC disease, cerebral toxoplasmosis, Parkinson's disease, multiple sclerosis,  
 CC brain cancers, hydrocephalus and encephalitis, and treat hepatic  
 CC disorders.

SO Sequence 3677 BP; 1055 A; 776 C; 881 G; 965 T; 0 other;

Alignment Scores:  
 Pred. No.: 12.8 Length: 3677  
 Score: 84.00 Matches: 55  
 Percent Similarity: 40.08% Conservative: 46  
 Best Local Similarity: 21.83% Mismatches: 75  
 Query Match: 7.51% Indels: 76  
 DB: 21 Gaps: 16

US-10-034-500-2 (1-218) x AAA75149 (1-3677)

OY 1 AAGlValThAlaSerCyThrIySArgValGluSerTyThrAsnIleuValAspTyr 20  
 DB 403 GCCCGAGACAGCCCTTCGAGAAAGACCTTGACCTCTAC--TTGCTCGAGACAG 459  
 OY 21 SerGlySerMetMetCysHisValAlaValArgGluProIleGluLeuAlaLys 40  
 DB 460 TCTGGAGTGTGGCAATAACGATTTGAATTATTAATTCGTACAGCAACTGGCGAG 519  
 OY 41 GluAlaIleLeuValIleAsnAlaAlaMetProIleMetSerTyGlnGlyLeuTyr 60  
 DB 520 AGATTGTG-----AGCCCGAATGAGATTATCTTATCTGTT 561  
 OY 61 ThrPheAlaProTyrSerValIleIlePro-----GlnGlySerTyrPheAsnSer 76  
 DB 562 TCT--TCTCAAGCACTATTATTGCTTACTGAGACAGAGCAAA-----609  
 OY 77 CyValAlaGluCysAlaValaIleAsnThrIleLysSerAspLeuGluIlePheGlyArgLeu 96  
 DB 610 -----ATCAGTAAAGCGCTTGAGAGATTAAACGTGTT 642  
 OY 97 ThrProValGly-----AspGlyIleLysMetHisGluThrValIleAsn 111  
 DB 643 AGTCGAGTAGAGAGACATATATCCATGAAGCACTAAAGCTCAATGCAATCAAAATTCAG 702  
 OY 112 Gln-----MetProProGlnAlaAlaValaIleLeuLeuThrAspGlyHisAsnAsn 128  
 DB 703 AAGCAGAGAGCGCTGAAACCTCCAGTATCAATTGCTTGACAGATGGCAAGTTGAC 762  
 OY 129 LeuGlyMetAsnPro-----ValGluGluValIleLys-----138  
 DB 763 ---GGTCTGCTGCATATATTCGAGAGAAAGGCAAAAGATATCCAGAGTTCAGTGGGCT 819  
 OY 139 SerIleTyGlnThrAsnProAsnValCysPheHisValaValSerPheAlaAspAspAla 158  
 DB 820 AGCTTTAT-----TGCTGTGGTCTCTGATTTT---GAAACAAGCA 858  
 OY 159 GluGlyValAlaIleLeuAsp-----GlnIleValaLeuAsnSerGlySerVal 175  
 DB 859 CAGCTTGAAGAATTCGTATTCACAGAGCAAGTTTCCCTTCGCAAGGTGAGTTTCAG 918  
 OY 176 LeuValAspGlyLeuGlnLeuGlnAsnProAlaValaCysGlnGluPheValAsnSer 195

DB 919 GCTCTTAAGAAATA-----ATTAATCT 942  
 OY 196 ValPheCysGln-----GluGlnIleLeuValThr 205  
 DB 943 ATACTAGCTCAGCTATGACTGAAATCTTCAAGTTCAGCCCTCAAGTGTCTGTGGG 1002  
 OY 206 GluGluVal---ValaValLeuArgGlyValaAsnPhe 216  
 DB 1003 GAGGAATTCAGATTGCTTAAAGTGAAGAGATTTC 1038  
 RESULT 35  
 ID AAA75157 standard; cDNA; 3677 BP.  
 AC AAA75157;  
 DT 15-JAN-2001 (first entry)  
 DE cDNA clone encoding a human TANGO 216 polypeptide.  
 KW TANGO 266; TANGO 216; TANGO 261; TANGO 262; TANGO 267;  
 KW cellular proliferation; cellular differentiation; cellular adhesion;  
 KW von Willebrand factor-associated disorder; cell trafficking; cancer;  
 KW hematopoietic associated disease; atelectasis; pulmonary congestion;  
 KW edema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;  
 KW intestinal disorder; spleen associated disease; renal disorder;  
 KW cardiovascular disorder; ischemic heart disease; hydrocephalus;  
 KW brain herniation; iatrogenic disease; inflammation; meningitis;  
 KW Alzheimer's disease; cerebral toxoplasmosis; Parkinson's disease;  
 KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT CDS 307..1773  
 FT /tag = a  
 FT /product = "TANGO 216"  
 PN WO200052022-A1.  
 PD 08-SEP-2000.  
 PF 01-MAR-2000; 2000WO-US05226.  
 PR 01-MAR-1999; 99US-0122458.  
 PA (MILL-) MILLENNIUM PHARM INC.  
 PI Barnes TM, Holtzman DA, Sharp JD, Fraser CC;  
 DR MPI: 2000-579269/54.  
 DR P-PSDB; AAB18455.  
 PT Novel human and murine secreted proteins designated TANGO 216, 261,  
 PT 262, 265 and 267 useful as modulating agents of cellular processes,  
 PT e.g. for treating cancer -  
 PS Disclosure; Page -, 175pp; English.  
 XX  
 CC AAA75157-59 encode human TANGO 216 proteins. The specification also  
 CC describes TANGO 266, TANGO 261, TANGO 262, and TANGO 267. The TANGO  
 CC polypeptides can be used to modulate cellular proliferation, modulate  
 CC cellular differentiation and/or modulate cellular adhesion. The  
 CC proteins can be used to treat any von Willebrand factor-associated  
 CC disorder, regulate extracellular matrix structuring, cellular adhesion,  
 CC and cell trafficking and/or migration, modulate cellular interactions,  
 CC modulate cell adhesion in proliferative disorders, such as cancer,  
 CC modulate the proliferation, differentiation, and/or function of cells  
 CC that appear in the bone marrow, and leukocytes, treat bone marrow, blood  
 CC and hematopoietic associated diseases and disorders, atelectasis,  
 CC pulmonary congestion or edema, emphysema, chronic bronchitis, bronchial  
 CC asthma and bronchiectasis, intestinal disorders, spleen associated







```

Oy 123 ThrAspGlyHisAsnAsnLeuGlyMetAsnProValGluGluValIleSerIleTyrGln 142
Db 2121 -----CAC-----TATCA 2113
Oy 143 ThrAsnProAsnValCys-----PheHisValVal 152
Db 2112 ACAGAACCGGCGACGTGCGGACATGTTTGTATTGATGTAATGTAATGTCCTG 2053
Oy 153 SerPheHisAspAspAlaIleGlyValAlaIleIleAspGlnIleVal-----AlaLeu 170
Db 2052 CAGCATGCTCAGAGATGCGCATGCAAGTGTCTCGGCGAGAGCGAGAGTGTAT 1993
Oy 171 AsnSerGlySerValIleuValAspGlyLeuGlnLeuGlnAsnProAlaValCysGln 190
Db 1992 GAAACTGCGCAAGTTGAGTGCATGTCAGTGCAGCCACAGGCTGAT-----CCGTCATGAG 1939
Oy 191 GluPhe-----ValAsnSerValPheCysGlnGluGlnIleLeuValThrGln 206
Db 1938 AAGTTTGGTACTGCTCATTCAGACACTTCATGTGTACATCATCATCTCATTTGCAAG 1879
Oy 207 GluValValValLeuArgGlyValAsnPhe 216
Db 1878 GTGGCTGCATCTACTGCGCTCATCTTACTTT 1849

RESULT 37
AAA81478
ID AAA81478 standard; DNA; 52253 BP.
AC AAA81478;
XX
XX 04-DEC-2000 (first entry)
DE N. meningitidis partial DNA sequence gnm_26 SEQ ID NO:26.
XX
XX Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
KM antigen; vaccine; diagnosis; infection; antibacterial; identification;
KM Meningococcus B; Menb; ds.
XX
XX Neisseria meningitidis.
OS
PN WO200022430-A2.
XX
XX 20-APR-2000.
PD
XX
XX 08-OCT-1999; 99WO-US23573.
PF
XX 09-OCT-1998; 98US-0103794.
PR 30-APR-1999; 99US-0132068.
XX
XX (CHIR ) CHIRON CORP.
PA
PI Fraser CM, Hickey E, Peterson J, Tettelin H, Venter JC,
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarcellini M, Scarlato V,
PI Rappelli R, Piazza M;
XX
XX MPI; 2000-318079/27.
DR
PT Isolated nucleotide sequences of Neisseria meningitidis which can be
PT used in the diagnosis and treatment of N. meningitidis infection and
PT other Neisserial infections, for example, N.gonorrhoea -
XX
XX Claim 7; Page 532-547; 1760pp; English.
XX
XX The present invention describes methods of obtaining immunogenic
XX proteins from Neisseria genomic sequences. AAA81453 to AAA82414
XX represent specifically claimed Neisseria meningitidis genomic DNA
XX sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
XX Neisseria DNA sequences and their corresponding proteins; AAA81254 to
XX AAA81255 and AAA81304 to AAA81321 represent PCR primers used in the
XX isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
XX AAA81452 represent Neisseria meningitidis Menb polynucleotide ORF
XX sequences, which are all used in the exemplification of the present

```

CC invention. The nucleic acid sequences, protein sequences, and antibodies  
CC against them, can be used in the manufacture of a composition. The  
CC composition can be used as a medicament (or in the manufacture of a  
CC medicament) for treating, preventing or diagnosing infection due to  
CC Neisserial bacteria. For example, some of the identified proteins could  
CC be components of vaccines against Meningococcus B; against all serotypes;  
CC and/or against all pathogenic Neisseriae. Identification of sequences  
CC from the bacterium will also facilitate production of biological probes,  
CC particularly organism-specific probes. Attempts to make efficacious  
CC Meningococcus B vaccines have failed mainly due to antigen tolerance.  
CC Multivalent vaccines have also been tried but none have successfully  
CC overcome antigenic variability. The provision of further, complete  
CC sequences may provide an opportunity to identify secreted or surface  
CC exposed proteins that may be presumed targets for the immune system and  
CC which are not antigenically variable or at least more conserved than  
CC other more variable regions.

XX Sequence 52253 BP; 12877 A; 15510 C; 13103 G; 10761 T; 2 other;

#### Alignment Scores:

Pred. No.:	624	Length:	52253
Score:	84.00	Matches:	37
Percent Similarity:	38.24%	Conservative:	28
Best Local Similarity:	21.76%	Mismatches:	58
Query Match:	7.51%	Indels:	47
DB:	21	Gaps:	5

US-10-034-500-2 (1-218) x AAA81478 (1-52253)

Oy 63 AlaProTyrSerValIleIleProGlnGlySerTPanSerCysValAlaGluCysAla 82

Db 47204 TCGCCTTCTTCAACGACGTTACCGCTGCTGTTAGATTTCATTTTCGGAATGCCGT 47263

Oy 83 Val-AsnThrIleIleYserSerAspLeuGluIlePheGlyArgLeuThrProValGlyAspG1 102

Db 47264 CTGAACCGCATTTCCGACACCAAGAACTGACATGAATGAAACAGATTGCTCCG 47323

Oy 102 YleIysMet----- 105

Db 47324 ACTGATGATGTTGGGATATGCGCTTTGGCACTTCCGACTTCACCATTCACAGATCCG 47383

Oy 106 -----HisGluThrValIleAsnGlnMetProProGlnAl 117

Db 47384 CGTCGAGGCTTGCAGCGACGACGCGAGCGAGTACGTTATTCACCTTCCGCCC----- 47435

Oy 117 AlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsnProValGluGluVal 137

Db 47436 -----GTCAAGTCGGGACACCTTACACACACACACAGCGAGTCCATC-----AT 47482

Oy 137 IlySerIleTyrGlnThrAsnProAsnValCysPheHisValIleSerPheAlaAspAs 157

Db 47483 CAAAACCTGTACGCCACC-----GGTTCTTTGACGA 47515

Oy 157 P-----AlaGluGlyValAlaIleIleAspGlnIleValAlaLeuAsnSe 172

Db 47516 CGTACCGCTCGAAACTGCGAGCGGACCTCTCGACCGTTATTCGAAAGCCGCCACCAT 47575

Oy 172 rGlySerValIleuValAspGlyLeuGlnLeuGlnAsnProAlaValCysGlnGluPuh 192

Db 47576 CGGCTGCTCAACATACCGGCGCAAAATGCTGCAAAACGAGCCATTAAAGAAAACCT 47635

Oy 192 eValAsnSerValPheCysGlnGluGln 201

Db 47636 CGAATGCTTGGGCTGCGGACATGCGCA 47663

#### RESULT 38

AAF21544/c

XX AAF21544 standard; DNA; 349980 BP.

XX AAF21544;

XX 13-MAR-2001 (first entry)

DE Neisseria meningitidis B nucleotide sequence SEQ ID NO:1.  
XX Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;  
KM diagnosis; antigen; detection; infection; gene therapy; antibacterial;  
KM de.  
XX Neisseria meningitidis.  
OS Neisseria meningitidis.  
PN WO20006791-A1.  
XX  
PD 09-NOV-2000.  
XX  
PF 08-MAR-2000; 2000WO-US05928.  
XX  
PR 30-APR-1999; 99US-0132068.  
PR 08-OCT-1999; 99WO-US23573.  
PR 28-FEB-2000; 2000GB-0004695.  
XX  
PA (CHIR) CHIRON CORP.  
PA (GENO-) INST GENOMIC RES.  
XX  
PI Plaza M, Hickey E, Peterson J, Tettelin H, Venter JC, Maignani V,  
PI Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V, Rappuoli R,  
PI Frazer CM, Grandi G;  
XX  
DR WPI; 2000-647603/62.  
XX  
PT Neisseria meningitidis B full length genome sequence and open reading  
PT frames are used to detect, treat and prevent Neisserial infections -  
XX  
PS Claim 7; Appendix A; 692pp; English.  
XX  
CC The present invention describes the full length genome of  
CC Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607  
CC to AAF21613 represent fragments of the NMB genomic sequence, as the  
CC sequence was too long to go in a record on its own it was split into 8  
CC sequences which overlap each other at the beginning and end of each  
CC sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at  
CC the beginning of AAF21607, the last 49980 bp of AAF21607 are repeated at  
CC the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the  
CC Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to  
CC AAF21606 represent PCR primers which are used in the exemplification of  
CC the present invention. The NMB genome and fragments from it have  
CC antibacterial activity, and can be used in vaccines and gene therapy.  
CC Neisseria nucleic acids, proteins and/or antibodies which binds to the  
CC proteins can be used in compositions for treating or preventing infection  
CC due to Neisserial bacteria or as a diagnostic reagent for detecting the  
CC presence of Neisserial bacteria or of antibodies raised to Neisserial  
CC bacteria. Computers, computer memory, computer storage medium or computer  
CC databases can be used in a search to identify open reading frames (ORFs)  
CC or coding sequences within the NMB genome. The DNA sequences provide  
CC further opportunities to find antigenic or immunogenic proteins which are  
CC more effective in vaccines than the outer membrane proteins currently  
CC used.  
XX  
SQ Sequence 349980 BP; 83241 A; 85091 C; 95206 G; 86442 T; 0 other;  
XX  
Alignment Scores:  
Pred. No.: 1.01e+04 Length: 349980  
Score: 84.00 Matches: 37  
Percent Similarity: 38.24% Conservative: 28  
Best Local Similarity: 21.76% Mismatches: 58  
Query Match: 7.51% Indels: 47  
DB: 21 Gaps: 5  
US-10-034-500-2 (1-218) x AAF21544 (1-349980)  
Qy 63 AAlaproyrServalIleIleProGInglySerTrpAasnSerCysValAlaGluCysAla 82  
:::|||||  
Db 180990 TCGCCTTCTTCAACAGCGTACCGCGCTCGCTGATTAGATTTCGATTCCGAATGCCGT 180841  
:::|||||  
Qy 83 ValAsnThrIleLeuSerAspLeuGluIlePheGlyArgLeuThrProValGlyAspG1 102  
:::|||||

Db 180840 CTGAACCGCATTCGCCGACCAAGAACTGACATGAAGAACTGATGCTTCGCC 180781  
Qy 102 yIleIySweT----- 105  
:::|||||  
Db 180780 ACTGATGATGTGGGCGATATGCGCTTGGACCTTGCCGACTTCACCATCAAGACATCGG 180721  
Qy 106 -----HIsGluThrValIleAnGlnMetProProGlnAl 117  
Db 180720 CGTCGAAGCTTGGACGCGTACCGAGCCGAGTACCGTATTCAACTACCTGCC----- 180669  
Qy 117 aAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsnProValGluGluVa 137  
:::|||||  
Db 180668 -----GTCAAGTCGCGCGACCTTACACACACACCGCGTGCATC-----AT 180622  
Qy 137 llySerIleTyroGlnThrAsnProAsnValCysPheHisValIleSerPheAlaAspAs 157  
:::|||||  
Db 180621 CAAGAAGCTGTACGCCAC-----GTTTCTTTGACGA 180589  
Qy 157 P-----AlaGluGlyValAlaIleIleAspGlnIleValAlaLeuAsnSe 172  
Db 180588 CGTACGCGTCGAACCTGCGAGCGGCGAGCTCTGCTGACCGTTATCGAACGCCACCAT 180529  
Qy 172 rGlySerValLeuValAspGlyLeuGlnLeuGlnAsnProAlaValCysGlnGluPh 192  
:::|||||  
Db 180528 CGGCTCGCTCAACATCACCGCGCAAAATCTGCAAAACGACCCATTAGAAAAAACC 180469  
Qy 192 eValAsnSerValPheCysGlnGluGln 201  
Db 180468 CGAATCGTTCGGCGTGGCGCGAGTCGCA 180441  
RESULT 39  
AAV59153  
ID AAV59153 standard; cDNA; 2573 BP.  
AC AAV59153;  
XX  
DT 14-DEC-1998 (first entry)  
XX  
DE Human ReproSA-1 coding sequence.  
XX  
KM 88; human; ReproSA-1; contraceptive; anti-sperm; fertility;  
KM in vitro fertilisation.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 125..1654  
FT /\*tag= a  
FT /product= "ReproSA-1 protein"  
XX  
PN WO9836073-A1.  
XX  
PD 20-AUG-1998.  
XX  
PF 17-FEB-1998; 98WO-US03513.  
XX  
PR 18-FEB-1997; 97US-0039574.  
XX  
PA (REPR-) REPROGEN INC.  
XX  
XX French CK, Neilson LJ, Schneider PA, Yamamoto KK;  
XX WPI; 1998-480771/41.  
XX P-PSDB; AAW77323.  
XX  
XX Sperm-specific polypeptide ReproSA-1 and related nucleic acid -  
XX transformed cells and antibodies, useful in contraceptive vaccines  
XX or compositions, and for diagnosis and treatment of infertility  
XX Claim 16; Fig 1; 62pp; English.  
XX  
XX The ReproSA-1 polypeptide, it's fragments, or nucleic acid encoding them  
XX are used in contraceptive vaccines to generate an anti-sperm response.  
CC

Antibodies can also be used as topical contraceptive (spermicide), also antisense sequences, ribozymes and triplex-forming molecules. Detection, by reaction with immobilised peptides, of anti-(II) antibodies (in serum, semen, saliva, cervical or vaginal mucosa) is used to diagnose immunological infertility and to monitor the effect of vaccination. Nucleotide fragments are used to diagnose infertility-associated mutations in the gene encoding Repro-SA, and for isolation of related sequences. Peptides and anti-idiotypic antibodies against the antibodies are used to inactivate the antibodies i.e. to increase fertility and may be added to sperm samples before in vitro fertilisation.

CC Sequence 2573 BP; 754 A; 532 C; 586 G; 701 T; 0 other;

Alignment Scores:  
Pred. No.: 8.75 Length: 2573  
Score: 83.50 Matches: 53  
Percent Similarity: 38.93% Conservative: 42  
Best Local Similarity: 21.72% Mismatches: 78  
Query Match: 7.46% Indels: 71  
DB: 19 Gaps: 14

US-10-034-500-2 (1-218) x AAV59153 (1-2573)

Oy 3 ValThrAlaSerCythrIysArgValGluSerTyraNtyrLeuValAspTyrsery 22  
Db 887 GTTGAACCTACTGCTGTGAAGACAAAGATGATAC-----GTGAAGAAAAATGCT 937  
Oy 23 SerMetMetLeuIshValAlaValArgGluProIle---GluLeuAlaIysGlu 41  
Db 938 TCTACTTATATAGAGATGCGAAACATACACCGACCTTCAACGCTGATTAAC 997  
Oy 42 Ala-----IleLeuIshLeuAlaIleMet 50  
Db 998 GCGAGAGGGGTGCTGCCGATTAAGCTGACGCTGCTGCAAGGAAACACGCGCTG 1057  
Oy 51 ProIysMetSerTyrsGlnIysGluLeuTythrPheAla-----ProTy 65  
Db 1058 CTTGGATCATGATGCTTGT-----TATGTAGCAGCTCATTCGAGAACCTAGCAATG 1111  
Oy 66 SerValIleIleProGlnIy-----SerTyrAsnSerCyValAlaGluCyValAla 83  
Db 1112 GCAGTCATCATTTCTAAGGCTGACCCAGTTGCTGCTGCTGACAGAAACCGGAA 1171  
Oy 84 AsnThrIleIysSer-----AspLeuGluIlePheGluArgLeuThrProVal 99  
Db 1172 GATCATATTAAGGCTGACGCTGCTGGGCTTGAAGACAGATTTGGAACACACTCTCT-- 1228  
Oy 100 GlyAspGlyIleIysMetHisGlu-----ThrValIleAsnGlnMetProProGln 116  
Db 1229 -----GAACACGACCGGCTGTTGCAGTCAAAATCTTGGCAATTCG 1273  
Oy 117 AlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGluMetAsnProValGluGlu 136  
Db 1274 CTTTCTTTGACATGTCACAGAAAGTTCTGAGGATCTCCAGTAAAGATAAAAGCC 1333  
Oy 137 ValIysSerIle-----Tyrgln 142  
Db 1334 ATTAAGAAATTCCTGCMAAAATGACTACTTACAGCCCTTGAACCATTTCTATATGAT 1393  
Oy 143 ThrAspProAsnValCyPheHisValIysSer-----PheAlaAspAsp 157  
Db 1394 GCTCCCTCCCAATTCGAAACATGTGTTGACAGCTTCAAGAGTCCCGCATGAT 1453  
Oy 158 AlaGluGlyIysAlaIleIleAspGlnIleValAlaLeuAsnSerGlySerValLeuVal 177  
Db 1454 AGCAAGCTCGACGACGATTTT-----GTAAACAAGTGT----- 1486  
Oy 178 AspGlyLeuGlnIleuLeuGlnAsn-----ProAlaValCySGlnIuphe 192  
Db 1487 ---GGCCTTTAAAAAGTTCAAGATAAAGACGAACTGGTCTTCCTTCAAGATAC 1543  
Oy 193 ValAsnSerVal 196

Db 1544 ATCAACGATATT 1555  
RESULT 40  
AAC39216  
ID AAC39216 standard, DNA, 2135 BP.  
XX  
XX AAC39216;  
XX  
XX 17-Oct-2000 (first entry)  
XX  
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 23804.  
DE  
XX  
XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; 88.  
XX Arabidopsis thaliana.  
XX  
XX EPI033405-A2.  
XX  
XX 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
XX 05-MAR-1999; 99US-0133180.  
XX 09-MAR-1999; 99US-0123548.  
XX 23-MAR-1999; 99US-0125788.  
XX 25-MAR-1999; 99US-0126264.  
XX 29-MAR-1999; 99US-0126785.  
XX 01-APR-1999; 99US-0127462.  
XX 06-APR-1999; 99US-0128234.  
XX 08-APR-1999; 99US-0128714.  
XX 16-APR-1999; 99US-0129845.  
XX 19-APR-1999; 99US-0130077.  
XX 21-APR-1999; 99US-0130449.  
XX 23-APR-1999; 99US-0130510.  
XX 23-APR-1999; 99US-0130891.  
XX 28-APR-1999; 99US-0131449.  
XX 30-APR-1999; 99US-0132048.  
XX 30-APR-1999; 99US-0132407.  
XX 04-MAY-1999; 99US-0132484.  
XX 05-MAY-1999; 99US-0132485.  
XX 06-MAY-1999; 99US-0132486.  
XX 06-MAY-1999; 99US-0132487.  
XX 07-MAY-1999; 99US-0132863.  
XX 11-MAY-1999; 99US-0134256.  
XX 14-MAY-1999; 99US-0134218.  
XX 14-MAY-1999; 99US-0134219.  
XX 14-MAY-1999; 99US-0134221.  
XX 14-MAY-1999; 99US-0134370.  
XX 18-MAY-1999; 99US-0134768.  
XX 19-MAY-1999; 99US-0134941.  
XX 20-MAY-1999; 99US-0135124.  
XX 21-MAY-1999; 99US-0135353.  
XX 24-MAY-1999; 99US-0135629.  
XX 25-MAY-1999; 99US-0136021.  
XX 27-MAY-1999; 99US-0136392.  
XX 28-MAY-1999; 99US-0136782.  
XX 01-JUN-1999; 99US-0137222.  
XX 03-JUN-1999; 99US-0137528.  
XX 04-JUN-1999; 99US-0137502.  
XX 07-JUN-1999; 99US-0137724.  
XX 08-JUN-1999; 99US-0138094.  
XX 10-JUN-1999; 99US-0138540.  
XX 10-JUN-1999; 99US-0138847.  
XX 14-JUN-1999; 99US-0139119.  
XX 16-JUN-1999; 99US-0139452.  
XX 16-JUN-1999; 99US-0139453.  
XX 17-JUN-1999; 99US-0139492.  
XX 18-JUN-1999; 99US-0139454.  
XX 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140635.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 26-JUL-1999; 99US-0145224.  
PR 27-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147933.  
PR 10-AUG-1999; 99US-0148319.  
PR 11-AUG-1999; 99US-0148317.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148688.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.

PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 13-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160771.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 26-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

## Alignment Scores:

Pred. No.:	7.68	Length:	2135
Score:	83.00	Matches:	57
Percent Similarity:	37.45%	Conservative:	40
Best Local Similarity:	22.01%	Mismatches:	88
Query Match:	7.42%	Indels:	74
DB:	21	Gaps:	14

US-10-034-500-2 (1-218) x AAC39216 (1-2135)

Cy 8 ThrlySArGValGIuSeTYrAenTYrleuValApTYrSerGIySer----- 23  
Db 1094 ACAAGAAGCGAGAAAGCCTTAACATCAGCAGAGATGTATCAGAGAGAGTGTTCACAA 1153

```
QY 24 -----MetMetMetLysHisValAlaValArgGluProLysIleGluLeuAlaLys 40
Db 1154 GCATTGCTTAAATGCTGGAAGGACAACTTGAATGTACTGAGAAAGCCGCTAGAAA 1213
QY 41 Glu----- 41
Db 1214 CACCCTAGAGTGACAAATATCCAGATAGACACAAAGACACTACTTCAATATGSGTGT 1273
QY 42 AlaIleLeuLysIleAsnAlaIleMetProLysMetSerTyrgInglyLeuTyrrhr 61
Db 1274 GCTTTTGTAGACATTGAAAAAGACTATCTCAGAGAGCGCCATGATTCTTCAATAGGTTT 1333
QY 62 PheAlaProTyrSerValIleIleProGlnGlySer---TrpAsnSerCysValAlaGlu 80
Db 1334 GGTCCTCCTGCTGACGCCAACATGAGGGCTGCTGCTTACAAACCCGCTGTGCATCA 1393
QY 81 CysAlaValAsnThrIleLys---SerAspLeuGlnIlePheGlyArgLeuThrProVal 99
Db 1394 AACCTAATGGAACCTGTGGAAGACAGTGAACCTGATTGCTTATGT---TTGATTCTGAA 1450
QY 100 GlyAspGly-----IleLysMetHisGluThrValIleAsnGlnMetPro 114
Db 1451 TTTGTTGACGATTCCAGTTCTGCTGATTTGCTTGTGACTGAGAACCGACTT--- 1507
QY 115 ProGlnAlaIleValIleLeuLeuThrAsnProAsnValHisAsnAsnLeuGlyMetAsnProVal 134
Db 1508 -----ATGCAGGCTGTGACAGAACCCAGAACCGCTTTGGA----- 1543
QY 135 GlnGluValLysSerIleTyrgInThrAsnProAsnValCysPheHisValIleSerPhe 154
Db 1544 AAGCAATACAAAAGATGTACAGATGAAT---AGTCTAAAGCTGCAT-----TTC 1591
QY 155 AlaAspAspAlaGlnGlyLysAlaIleIleAspGlnIleValAlaLeuAsnSerGly--- 173
Db 1592 ACAGAAAGCGCACTG---CGCTTATAGCTTGAAGAAAGCAATTAAGAAAACACGGGTGCT 1648
QY 174 -----SerValIleuValAspGlyLeuGlnLeuGlnAsn 185
Db 1649 CGTGGCTTAAGACTCTTCTAGAAAGCATCTCATGACTATGTACGAGATTCCGAGAC 1708
QY 186 ProAlaValCysGlnGlnPheValAsnSerValPheCysGlnGlnGlnIleuValThr 205
Db 1709 GAGGATAACAGCAGCATATGATA-----GAAAGCGGTTGTGTGTGAT 1750
QY 206 GlnGluValVal-----ValLeuArgGly 213
Db 1751 GAAGAAACAGTTGAAGAGAAAGGCGCTGAGGCTTGGAGCTAAGATTCTCCGTGT 1807
```

Search completed: April 8, 2003, 05:25:08  
Job time : 1513 secs

GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 8, 2003, 05:01:54 ; Search time 49 Seconds  
(without alignments)  
1364.399 Million cell updates/sec

Title: US-10-034-500-2  
Perfect score: 1119  
Sequence: 1 AEWASCTKVESNYLVY.....QEQILVTEVVLRGVNFAF 218

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 6.0 , Ygapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues  
Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODL=frame+ p2n model -DEV=xlh  
-Q=/cgn2\_1/USPFO\_epool/US10034500/runat\_02042003\_092745\_19735/app\_query.fasta\_1.391  
-DB=Issued\_Patents\_NA -QWrt=faaetap -SUFFIX=rml -MINMATCH=0.1 -LOOPCL=0  
-LOOEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=b10sum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pcr -THR MAX=100 -THR MIN=0 -ALIGN=40  
-MODE=LOCAL -OUTFMT=ptco -NORW=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10034500 @CGN 1 1 32 @runat\_02042003\_092745\_19735 -NCPU=6 -ICPU=3  
-NO\_XLPRY -NO\_MMAR -LARGQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WART TIMEOUT=30 -THRAD=1 -XGAPOP=10 -XGAPEXT=0.5 -FAPAPOP=6 -FAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA: \*  
1: /cgn2\_6/ptodata/1/ina/5A COMB.seq: \*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq: \*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq: \*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq: \*  
5: /cgn2\_6/ptodata/1/ina/PTUS COMB.seq: \*  
6: /cgn2\_6/ptodata/1/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	91.5	8.2	3043	4 US-09-049-698-16	Sequence 16, Appl
2	91.5	8.2	3181	4 US-09-049-698-18	Sequence 19, Appl
3	86	7.7	3668	4 US-09-206-942-36	Sequence 36, Appl
4	86	7.7	3686	4 US-09-206-942-33	Sequence 33, Appl
5	83.5	7.5	2556	4 US-09-221-017B-874	Sequence 874, App
6	81	7.2	3250	4 US-09-221-017B-960	Sequence 960, App
7	80.5	7.2	1380	3 US-08-968-563-5	Sequence 5, Appli
8	80.5	7.2	1380	3 US-08-968-563A-5	Sequence 5, Appli
9	80.5	7.2	1380	4 US-09-297-928-1	Sequence 1, Appli
10	79.5	7.1	3288	4 US-09-206-942-44	Sequence 44, Appl
11	79.5	7.1	3306	4 US-09-206-942-42	Sequence 42, Appl
12	79	7.1	423	1 US-08-470-179-89	Sequence 89, Appl

C 13	79	7.1	3789	1 US-07-872-644-42	Sequence 42, Appl
C 14	79	7.1	3789	1 US-08-297-494-42	Sequence 42, Appl
C 15	79	7.1	3789	1 US-08-297-510-42	Sequence 42, Appl
C 16	79	7.1	3789	1 US-08-479-532-42	Sequence 42, Appl
C 17	79	7.1	3789	1 US-08-455-526-42	Sequence 42, Appl
C 18	79	7.1	3789	1 US-08-455-525-42	Sequence 42, Appl
C 19	79	7.1	3789	5 US-09-139-491-42	Sequence 42, Appl
C 20	79	7.1	3789	5 PCT-US92-03222-42	Sequence 42, Appl
C 21	79	7.1	4131	1 US-07-872-644-38	Sequence 38, Appl
C 22	79	7.1	4131	1 US-08-297-494-38	Sequence 38, Appl
C 23	79	7.1	4131	1 US-08-297-510-38	Sequence 38, Appl
C 24	79	7.1	4131	1 US-08-479-532-38	Sequence 38, Appl
C 25	79	7.1	4131	1 US-08-455-526-38	Sequence 38, Appl
C 26	79	7.1	4131	1 US-08-455-525-38	Sequence 38, Appl
C 27	79	7.1	4131	3 US-09-139-491-38	Sequence 38, Appl
C 28	79	7.1	4131	5 PCT-US92-03222-38	Sequence 38, Appl
C 29	79	7.1	15894	1 US-08-348-891A-1	Sequence 1, Appli
C 30	79	7.1	15894	1 US-08-905-817-1	Sequence 1, Appli
C 31	77.5	6.9	4403765	4 US-09-103-840A-2	Sequence 2, Appli
C 32	77	6.9	423	1 US-08-470-179-93	Sequence 93, Appl
C 33	77	6.9	960	2 US-08-245-511-3	Sequence 3, Appli
C 34	77	6.9	960	2 US-08-600-993A-3	Sequence 3, Appli
C 35	77	6.9	1999	3 US-08-961-083-1	Sequence 1, Appli
C 36	77	6.9	2049	3 US-08-481-435-5	Sequence 5, Appli
C 37	77	6.9	10711	4 US-08-961-527-145	Sequence 145, App
C 38	76.5	6.8	2178	4 US-08-936-165A-199	Sequence 199, Appl
C 39	76.5	6.8	3007	4 US-09-193-562D-27	Sequence 27, Appl
C 40	76.5	6.8	3839	4 US-09-245-248B-54	Sequence 54, Appl
C 41	76.5	6.8	4376	1 US-08-119-125A-1	Sequence 1, Appli
C 42	76.5	6.8	4965	1 US-08-961-527-143	Sequence 143, App
C 43	76.5	6.8	6744	1 US-08-119-125A-2	Sequence 2, Appli
C 44	76.5	6.8	20986	4 US-08-961-527-54	Sequence 54, Appl
C 45	76	6.8	1992	1 US-08-485-284A-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1  
US-09-049-698-16  
Sequence 16, Application US/09049698  
Patent No. 6368792  
GENERAL INFORMATION:  
APPLICANT: BILLING-MEDEL, PATRICIA A.  
APPLICANT: COHEN, MAURICE  
APPLICANT: COLPITTS, TRACEY L.  
APPLICANT: FRIEDMAN, PAULA N.  
APPLICANT: HAYDEN, MARK  
APPLICANT: KLAAS, MICHAEL R.  
APPLICANT: ROBERTS-RAPP, LISA  
APPLICANT: RUSSELL, JOHN C.  
APPLICANT: STROUPE, STEPHEN D.  
TITLE OF INVENTION: REAGENTS AND METHODS FOR THE  
TITLE OF INVENTION: USEFUL FOR DETECTING DISEASES OF THE GASTROINTESTINAL  
NUMBER OF INVENTION: TRACT  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESS:  
ADDRESS: Abbott Laboratories  
STREET: 100 Abbott Park Road  
CITY: Abbott Park  
STATE: IL  
COUNTRY: USA  
ZIP: 60064-3500  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/049,698  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/828,856

```

FILING DATE: 31-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6068.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 3043 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-049-698-16

Alignment Scores:
Pred. No.: 0.0308 Length: 3043
Score: 91.50 Matches: 46
Percent Similarity: 35.64% Conservative: 26
Best Local Similarity: 22.77% Mismatches: 55
Query Match: 8.18% Indels: 75
DB: Gaps: 8

US-10-034-500-2 (1-218) x US-09-049-698-16 (1-3043)
Qy 17 LeuValAspTyrSerGlySerMetMetMetLeuHisValAlaValArgIuProlysile 36
Db 941 GTTCTGATAGTCTGGAGACATGGG----- 967
Qy 37 GluLeuAlaValGluAlaIleLeuLysIleAsnAlaIleMetProLysMetSerTyrGln 56
Db 968 -----GGTAAAGGACCCGCTAAATCGAATATCAACAGCAGCAAAACATTTCCTGCTGAG 1021
Qy 57 GlyGlyLeuTyrThrPheAlaProTyrSerValIleIleProGlnIleSerThrPasnSer 76
Db 1022 -----ACTGTGAAAATGATCTCGGTGGGG 1048
Qy 77 CysVal-----AlaGlu 80
Db 1049 ATGGTTCACTTGTGATGACTGACCACTATTGTAAATAGCTAATCAATAAAGACAGT 1108
Qy 81 CysAlaValAsnThrIleLysSerAspLeuGlnIlePheGlyArgLeuThrProValGly 100
Db 1109 GATGAAGAAACACACTCATGGCAGATTCTCAACATAT-----CCTCGGGA 1156
Qy 101 -----AspGlyIleLysMetHisGluThrValIleAsnGlnMetProPro 115
Db 1157 GGAACCTTCATCTGCTCTGGAAATTAAATATGACATTTCAGGTGATTGGAAGCTACATTCC 1216
Qy 116 Gln-----AlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlnMetAsn 132
Db 1217 CAATCGATGATGATCGGAAGTACTCTCTGACTGATGGGAGGATAAACATCGCAAGTTCT 1276
Qy 133 ProValGluGluValLysSerIleTyrGlnThrAsnProAsnValCysPheHisValVal 152
Db 1277 TGTATGATGAAGTGAAG-----CAAGTGGGGCCATTGTTCATTTTATGTCTTGG 1322
Qy 153 SerPheAla----- 155
Db 1328 GGAAGACCTGCTGATGAAGCAGTAAATAGAGATGAGCAAGATAACAGAGAAATCATTTT 1387
Qy 156 -----AspAspAlaGluGlyLysAlaIleIleAspGlnIleValAlaLeuAsnSer 172
Db 1388 TATGTTTCAGATGAAGCTCAGAACAAATGCGCTCATTGATCTTTTGGGGCTCTTACATCA 1447
Qy 173 GlySer 174
Db 1448 GGAAT 1453

RESULT 2
US-09-049-698-18

```

```

Sequence 18, Application US/09049638
Patent No. 6368792
GENERAL INFORMATION:
APPLICANT: BILLING-MBEDEL, PATRICIA A.
APPLICANT: COHEN, MADRICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: HAYDEN, MARK
APPLICANT: KLAAS, MICHAEL R.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS FOR THE
TITLE OF INVENTION: USEFUL FOR DETECTING DISEASES OF THE GASTROINTESTINAL
TITLE OF INVENTION: TRACT
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/049,698
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/828,856
FILING DATE: 31-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6068 US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 3181 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-049-698-18
Alignment Scores:
Pred. No.: 0.0331 Length: 3181
Score: 91.50 Matches: 46
Best Local Similarity: 35.64% Conservative: 26
Query Match: 22.77% Mismatches: 55
DB: 8.18% Indels: 75
Gaps: 8
US-10-034-500-2 (1-218) x US-09-049-698-18 (1-3181)
Qy 17 LeuValAlaPtyrSerGlySerMetMetMetClyHisValAlaValArgGluProLysIle 36
Db 952 GTTCTGTAAGTACTGCGAAGCATGGG----- 978
Qy 37 GluLeuAlaLysGluAlaIleLeuLysIleAlaAlaIleMetProLysMetSerTyrGln 56
Db 979 -----GGTAAGACCGCGCTAAATGCAATGCAATCAAGCAAAACATTTCCTGCTGCAG 1032
Qy 57 GlyGlyLeuTyrThrPheAlaProLysSerValIleIleIleProGlnGlySerThrPasnser 76
Db 1033 -----ACTGTTGAAATGAGATCCTGGGATGGGG 1059

```

Qy 77 CybVal1-----AlaGlu 80  
Db 1060 ATGGTCACTTGTAGTAGTACGCACTATTGTAAATAAGCTAATCAATAAAGCAGT 1119  
Qy 81 CybAlaValAanthr11leuSerAspLeuGlu1lePheGlyArgLeuThrProValGly 100  
Db 1120 GATGAAAGAACACACTGAGCAGATTAACCTACATAT-----CCTCTGGGA 1167  
Qy 101 -----AspGly1leuSerMetHisGluThrVal1leAsnGlnMetProPro 115  
Db 1168 GGAACCTTCATCTGCTCGGATTAATATGATTCACGTATTCAGAGCTACATTC 1227  
Qy 116 Gln-----AlaAlaVal1leuLeuThrAspGlyHisAsnAsnLeuGlyMetAsn 132  
Db 1228 CAATCGATGATGCGAAGTACTGCTGCTGATGATGGAGAGATTAACCTCAAGTTCT 1287  
Qy 133 ProValGluGluValValSer1leuThrGlnThrAsnProAsnValCysPheHisValVal 152  
Db 1288 TGTATTGATGAAGTAAA-----CAAGTGGGCGCATTTTCATTTATTCCTTTG 1338  
Qy 153 SerPheAla----- 155  
Db 1339 GGAAGAGCTGCGTAGAAGCACTAATAGATGAGCAAGATTAACAGAGAGTCAATTT 1398  
Qy 156 -----AspAspAlaGluGlyValAla1le1leAspGln1leValAlaLeuAsnSer 172  
Db 1399 TATGTTTCAATGAGAGTCAAGAACATGCGCTCATTTGATGCTTTGGGCTCTTACATCA 1458  
Qy 173 GlySer 174  
Db 1459 GGAAT 1464  
RESULT 3  
US-09-206-942-36  
; Sequence 36, Application US/09206942  
; Patent No. 6432869  
; GENERAL INFORMATION:  
; APPLICANT: Looemore, Sheena M.  
; APPLICANT: Yang, Yan-Ping  
; APPLICANT: Klein, Michel H.  
; TITLE OF INVENTION: Protective Recombinant Haemophilus Influenzae High  
; TITLE OF INVENTION: Molecular Weight Proteins  
; FILE REFERENCE: 1038-861 MIS:jb  
; CURRENT APPLICATION NUMBER: US/09/206,942  
; EARLIER FILING DATE: 1998-12-08  
; EARLIER APPLICATION NUMBER: 09/167,568  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 36  
; LENGTH: 3668  
; TYPE: DNA  
; ORGANISM: Haemophilus influenzae  
US-09-206-942-36  
Alignment Scores:  
Pred. No.: 0.246 Length: 3668  
Score: 86.00 Matches: 48  
Percent Similarity: 31.85% Conservative: 38  
Best Local Similarity: 17.78% Mismatches: 96  
Query Match: 7.69% Indels: 86  
Gaps: 7  
US-10-034-500-2 (1-218) x US-09-206-942-36 (1-3668)  
Qy 11 ValGluSer1TyRAsn1TyRLeuValAsp1TySerGlySerMetMetCelyHisValAla 30  
Db 984 ATAGAGTCAACAAACTTTAGTCCCTCAGAGAGGTCAAGCTTAATAATTCAAAAGCAAGGT 1043  
Qy 31 ValArgGluProLeu1leGluLeuAlaLysGluAla1leLeu----- 44  
Db 1044 TCGACACAGCGCGCTTTTACATTAATAAATAATGATTAAATTAAATGCCACTGGGGCAAT 1103

Qy 45 -----Lys1leAsnAla1leMet 50  
Db 1104 ATATCATGAACCAAGTTCAGATTTAGTATGATATCTCAAAAAAGCCTTATAGCCAAAT 1163  
Qy 51 ProLysMetSer1TyR1Gln1Gly1leu1TyR1PheAla----- 63  
Db 1164 AAAACATACACTTTGAAAGGGGCAATATCACTCTTGACGCCGATTAATAAACCAATAGAA 1223  
Qy 64 ---Pro1TySerVal1le1leProGln1GlySer1TyRAsnSerCysValAlaGluCysAla 82  
Db 1224 ATCAAGATATATATTCATCTTTAAGAGAGAGCC---AATGTCACCCCTTCGATAGCCGAAT 1280  
Qy 83 ValAsnThr1leLysSerAspLeuGlu1lePheGlyArgLeuThrProValGlyAspGly 102  
Db 1281 TATGTAATAGCAATCAAGCTTTAAGTTAAGAGAAATGACATTAATTAAGGCAAT--- 1337  
Qy 103 1leuSerMetHisGluThrVal1leAsnGlnMetProProGln1leAlaVal1leLeuLeu 122  
Db 1338 -----CTCACCGTT 1346  
Qy 123 ThrAspGlyHisAsnAsnLeuGlyMetAsnProValGluGluValLysSer1leu1TyR1Gln 142  
Db 1347 ACCGGCTCCGCTATCATATATGAAAAAATCTTACCTTGAAAGTAGTACATTTT 1406  
Qy 143 ThrAsnProAsnValCysPheHisValValSerPheAlaAspAsp----- 157  
Db 1407 GCTATTCAAATTAACGCTTTAAGCTATCCGCTATTTTGACCAACAGCAAGCAATCAAC 1466  
Qy 158 -----AlaGluGlyValAla1le1leAspGln1leValAlaLeuAsnSerGlySer 174  
Db 1467 ATTTCCATCGCTTAAGGAGGAGCATTTTAAAGATATCGAG-----AATATGGCAGT 1520  
Qy 175 ValLeuVal----- 177  
Db 1521 CTGAATATTAACCTAATTCGAGCTCCACCAACCACTATTATTAAGGTAAATTAAT 1580  
Qy 178 -----AspGlyLeuGlnLeuGlnAsn 185  
Db 1581 AACGAAAGAGTATTAAATATACGAAATATGATGATTAATACGAATCAAAATGTGGC 1640  
Qy 186 ProAlaValCysGlnGluPheValAsnSerValPheCysGlnGluGln1leLeuValThr 205  
Db 1641 GGCATATATCTCGCAAAAAGAGGCAATCTCAAAATTTCTGTATTAAGTCAATATTACC 1700  
Qy 206 GluGluValValValLeuArg1TyValAsn 215  
Db 1701 GAGCGATTAACATCAAGCAGCGTTAAT 1730  
RESULT 4  
US-09-206-942-33  
; Sequence 33, Application US/09206942  
; Patent No. 6432869  
; GENERAL INFORMATION:  
; APPLICANT: Looemore, Sheena M.  
; APPLICANT: Yang, Yan-Ping  
; APPLICANT: Klein, Michel H.  
; TITLE OF INVENTION: Protective Recombinant Haemophilus Influenzae High  
; TITLE OF INVENTION: Molecular Weight Proteins  
; FILE REFERENCE: 1038-861 MIS:jb  
; CURRENT APPLICATION NUMBER: US/09/206,942  
; EARLIER FILING DATE: 1998-12-08  
; EARLIER APPLICATION NUMBER: 09/167,568  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 33  
; LENGTH: 3686  
; TYPE: DNA  
; ORGANISM: Haemophilus influenzae  
US-09-206-942-33  
Alignment Scores:  
Pred. No.: 0.248 Length: 3686



Score: 86.00 Matches: 48  
 Percent Similarity: 31.85% Conservative: 38  
 Best Local Similarity: 17.78% Mismatches: 98  
 Query Match: 7.69% Indels: 86  
 DB: 4 Gaps: 7

US-10-034-500-2 (1-218) x US-09-206-942-33 (1-3686)

```

QY 11 ValGluSerTyrAspTyrLeuValAspTyrSerGlySerMetMetLeuValAla 30
DB 1002 ATAGAGTCAAAATCTTGTGCTCAGAGGCTCAAGCTTAAATTCAAAGCCAAAGT 1061
QY 31 ValArgGluProLysIleGluLeuAlaLysGluAlaIleLeu 44
DB 1062 TCGACACAGCGCGCTTTACATTAATAATTTAAATTCAGCTGGGGCAAT 1121
QY 45 -----LysIleAsnAlaIleLeu 50
DB 1122 ATATGATGACCAAGTTGACAGGTATGATGATATCTCAAAAAAGCCTTAGCCAAAT 1181
QY 51 ProLysMetSerTyrGlnGlyLeuTyrThrPheAla 63
DB 1182 AAAACATACCTTTCGAGGGCAATATACCTTTCAGCCGATTAATAAACCAATAGA 1241
QY 64 ---ProTyrSerValIleIleProGlnGlySerTyrPheSerCysValAlaGluCysAla 82
DB 1242 ATCAAAAGTAATATCTCTGTAAAGAGAGCC---AATGTCACTTGTGTCAGCGCAAT 1298
QY 83 ValAsnThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGlyAspGly 102
DB 1299 TATGTATATGACCAATACCTTAACTAAGAGAAAGTCACTAATAAGGCAT--- 1355
QY 103 IleLysMetIleGluThrValIleAsnGlnMetProProGlnAlaIleValIleLeuLeu 122
DB 1356 -----CTCACCGTT 1364
QY 123 ThrAspGlyHisAsnAsnLeuGlyMetAsnProValGluGluValLysSerIleTyrGln 142
DB 1365 ACCGGCTCCGCTTCAATATAGAAAAAATCTTACCGTTGAGGTAGTGAAGTTTAA 1424
QY 143 ThrAsnProAsnValCysPheHisValValSerPheAlaAsp 157
DB 1425 GCTAATCTCAATTAACGCTTAACTGATCCGCTATTGACACCAAGCAAGTCAAAAC 1484
QY 158 -----AlaGluGlyLysAlaIleIleAspGlnIleValAlaLeuAsnSerGlySer 174
DB 1485 ATTTCCATGCTTAAGGAGGAGCTATTTTAAAGATATGAG---AATACTGGCAAT 1538
QY 175 ValLeuVal----- 177
DB 1539 CTGAATATTAACCACTAATTCGACTCCAAACCACTACTATTATAAGGTAATATACT 1598
QY 178 -----AspGlyLeuGlnLeuLeuGlnAsn 185
DB 1599 AACAGAAAAGTATTAATATATACGAATATATGCTGATATATCTGAATCCAAATTCGC 1658
QY 186 ProAlaValCysGlnGluPheValAsnSerValPheCysGlnGluGlnIleLeuValThr 205
DB 1659 GGCATATCTCGCAAAAGAGCAATCTCACAATTTCTTCTGATTAAGTCATATTATACC 1718
QY 206 GluGluValValValLeuArgGlyValAsn 215
DB 1719 GAGCGGATTAACATCAAGCAAGCGGTTAAT 1748

```

```

ADDRESS: MORRISON & FOERSTER
STREET: 755 PAGE MILL ROAD
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/221, 017B
FILING DATE: 23-DEC-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP1182
FILING DATE: 31-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP1546
FILING DATE: 30-JAN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP2911
FILING DATE: 09-APR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU98/01023
FILING DATE: 10-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Monroy, Gladys H
REGISTRATION NUMBER: 32,430
REFERENCE/DOCKET NUMBER: 27340-20021.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-813-5600
TELEX: 650-494-0792
INFORMATION FOR SEQ ID NO: 874:
SEQUENCE CHARACTERISTICS:
LENGTH: 256 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: UNKNOWN
ORIGINAL SOURCE:
ORGANISM: PORYPHYROMONAS GINGIVALIS
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1...2556
US-09-221-017B-874
Alignment Scores:
Pred. No.: 0.305 Length: 2556
Score: 83.50 Matches: 36
Percent Similarity: 44.88% Conservative: 21
Best Local Similarity: 28.35% Mismatches: 37
Query Match: 7.46% Indels: 33
DB: 4 Gaps: 8
US-10-034-500-2 (1-218) x US-09-221-017B-874 (1-2556)
QY 18 ValAspTyrSerGlySerMetMetLeuValAlaValArgGluProLysIleGln 37
DB 750 ATGATGTCGACGAGCT-----ATGCAAGGCAATGCAATTCGCAAGCAATGCTGAG 803
QY 38 LeuAlaLysGluAlaIleLeuLysIleAsnAlaIleMetProLysMetSerTyrGlnGly 57
DB 804 GCAAGCAAGAGTACGATTCCTTATCAACAATTCGTCGACAGCAATATTC--- 857
QY 58 GlyLeuTyrThrPheAlaProTyrSerValIleIleProGlnGlySerTyrPheSerCys 77
DB 858 GGAATGTTACATTCGCGGAGAGAGC----- 884

```



```

Oy 44 LeuysIIeaNaIaIaMeProIysmeSerYrGInglYleuYrYrIhPheAla 63
Db 97 ATAAAGATCTGCTGCTGCTGATGATTAATTAATTAATTCGGCACCTGAATGCTGGT 156
Oy 64 -----ProYSerVal--- 67
Db 157 AGAAAGAGAGTCTCTTCTGTTCTTTGAAGCTCGCGAAAGACCTTCAAGGTACT 216
Oy 68 ILeIleProGInglYserTTrpAsnSerCySAlaIaGInCySAlaValaLenTrIleYe 87
Db 217 GGTATGGAATCTGGTAATCTGGAGGATCTAATTAATGCAAGGTGGTGGCCGAAATTTGTAAG 276
Oy 88 SerAspLeuGInIlePheGlyArgIleuThxProValGly-----Asp 101
Db 277 GGATGCCCAAGAAAGTTTGGCTGCATATAGTACAATGTGGGTGTCGAAGAAAGATCAAT 336
Oy 102 GlyIleYsMetHisGluThVal-----IleAsnGInMetProPro 115
Db 337 GGTGAATAATTAATCTGAATCATTAATTAATCTAGACATCAAAAGGAATATCTGGCGC 386
Oy 116 GInIaIaIaValIleLeuLeuThAspGlyHisAsnIleuGInMetAsnPro----- 133
Db 397 -----ATGACTCTACCCGAC-----AATTGGTGGCTAAATCCAGACTTG 425
Oy 134 ValGInGluValYsSerIle-----TyGInTrhAsnPro 145
Db 436 ATTGATGATGATCAAGATGTGCATCATCATCGTTTTCACATTTCCACATCAATTTTGGCC 495
Oy 146 AsnValCYsrPheHisValValSerTherhAlaSerhAlaGInGlyYsAlaIleIleAsp 155
Db 436 GCTATCTGTAGCCCAATTGAAGATCATGTAT-----TCCACGTGAGA 540

```

[illegible]

```

Db 97 ATAAAGATGCTGCTGCTGCTGATGATTAACCTTAACCTTCCGCCACTTGAACTGCT 156
Qy 64 -----ProTyrSerVal--- 67
Db 157 AGAAGAGAGAGTCTCTCTCTCTCTTTTGAAGGCTGCCGAAAGACCTTTCAGAGTTACT 216
Qy 68 ILeIleProGInglySerTyrPasnSerCyValAlaGluCyValAlaValAsnThrIleLys 67
Db 217 GTGATTGATCTGGTGAAGGGGTACTACTATTTGCCAAGGTGTTGCCGAAATTTGTAAG 276
Qy 88 SerAspLeuGluIlePheGlyArgLeuThrProValGly-----Asp 101
Db 277 GGAATACCAAGAGTTTGTCTCCATGATACAAATGGGTGTTGGAAGAGATCAAT 336
Qy 102 GlyIleLysMetHisGluThrVal-----IleAsnGlnMetProPro 115
Db 337 GGTGAAAATTGACTGAAATCATTAATCTAGACATCAAAACGTGAATATTGCTGCTGCC 396
Qy 116 GlnAlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsnPro----- 133
Db 397 -----ATCAGCTTACCCGAC-----AATTGGTTGCTTAATCCAGACTTG 435
Qy 134 ValGluGluValValysSerIle-----TyrGlnThrAsnPro 145
Db 436 ATGTGATTCAGTCAAGAGATGTCAGATCATCGTTTCAACATTCACATCAATTTTGGCC 495
Qy 146 AsnValCySPheHisValIleSerPheAlaAspAspAlaGluGlyValAlaIleLeuAsp 165
Db 496 CGTATCTGTACGCAATTGAAAGGTGATGTTGAT-----TCACAGCTCAGA 540
Qy 166 GlnIleValAlaLeuAsnSerGlySerValLeuValAspGlyLeuGlnLeuGlnAsn 135
Db 541 GCTATCTCTGCTTAAGGGTTTGAAGTGGTGTAAAGGTGTCATTTGCTATCTCT 600
Qy 186 ProAlaValCySGInGluPheValAsnSerValPheCySGInGluGlnIleLeuValThr 205
Db 601 -----TACATCACT 609
Qy 206 GlnGluValValVal-----LeuArgGlyValAlaAsnPheAla 217
Db 610 GAGGAAGTACGATTCATCAATGTGCTCTATGTGGTGAACATTTGCC 657

```

RESULT 9  
 US-09-297-928-1  
 Sequence 1, Application US/09297928  
 Patent No. 6358716  
 GENERAL INFORMATION:  
 APPLICANT: BULTHUIS, BEN A.  
 GATENBY, ANTHONY A.  
 HAYNIE, SHARON L.  
 HSU, AMY K.  
 LARBEAU, RICHARD D.  
 TITLE OF INVENTION: METHOD FOR THE PRODUCTION OF  
 GLYCEROL BY RECOMBINANT  
 ORGANISMS  
 NUMBER OF SEQUENCES: 25  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: GENENCOR INTERNATIONAL, INC.  
 STREET: 4 CAMBRIDGE PLACE  
 1870 SOUTH WINTON ROAD  
 CITY: ROCHESTER  
 STATE: NEW YORK  
 COUNTRY: U.S.A.  
 ZIP: 14618  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: DISKETTE, 3.5 INCH  
 COMPUTER: IBM PC COMPATIBLE  
 OPERATING SYSTEM: MICROSOFT WINDOWS 95  
 SOFTWARE: MICROSOFT WORD VERSION 7.0A  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/297,928  
 FILING DATE: 11-May-1999

```

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/03602
FILING DATE: NOVEMBER 13, 1996
ATTORNEY/AGENT INFORMATION:
NAME: FLOYD, LINDA AXAMETHY
REGISTRATION NUMBER: 33,692
REFERENCE/DOCKET NUMBER: CR-9981-P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-892-8112
TELEFAX: 302-773-0164
TELEX: 6717325
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1380 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-297-928-1

Alignment Scores:
Pred. No.: 0.294 Length: 1380
Score: 80.50 Matches: 44
Percent Similarity: 35.65% Conservative: 33
Best Local Similarity: 20.37% Mismatches: 68
Query Match: 7.19% Indels: 71
DB: Gaps: 11

US-10-034-500-2 (1-218) x US-09-297-928-1 (1-1380)
Qy 44 LeuYsIleAsnAlaValMetProLysMetSerTyrGInglyLeuTyrThrPheAla 63
Db 97 ATAAAGATGCTGCTGCTGCTGATGATTAACCTTAACCTTCCGCCACTTGAACTGCT 156
Qy 64 -----ProTyrSerVal--- 67
Db 157 AGAAGAGAGAGTCTCTCTCTCTCTTTTGAAGGCTGCCGAAAGACCTTTCAGAGTTACT 216
Qy 68 ILeIleProGInglySerTyrPasnSerCyValAlaGluCyValAlaValAsnThrIleLys 87
Db 217 GTGATTGATCTGGTGAAGGGGTACTACTATTTGCCAAGGTGTTGCCGAAATTTGTAAG 276
Qy 88 SerAspLeuGluIlePheGlyArgLeuThrProValGly-----Asp 101
Db 277 GGAATACCAAGAGTTTGTCTCCATGATACAAATGGGTGTTGGAAGAGATCAAT 336
Qy 102 GlyIleLysMetHisGluThrVal-----IleAsnGlnMetProPro 115
Db 337 GGTGAAAATTGACTGAAATCATTAATCTAGACATCAAAACGTGAATATTGCTGCTGCC 396
Qy 116 GlnAlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsnPro----- 133
Db 397 -----ATCAGCTTACCCGAC-----AATTGGTTGCTTAATCCAGACTTG 435
Qy 134 ValGluGluValValysSerIle-----TyrGlnThrAsnPro 145
Db 436 ATGTGATTCAGTCAAGAGATGTCAGATCATCGTTTCAACATTCACATCAATTTTGGCC 495
Qy 146 AsnValCySPheHisValIleSerPheAlaAspAspAlaGluGlyValAlaIleLeuAsp 165
Db 496 CGTATCTGTACGCAATTGAAAGGTGATGTTGAT-----TCACAGCTCAGA 540
Qy 166 GlnIleValAlaLeuAsnSerGlySerValLeuValAspGlyLeuGlnLeuGlnAsn 185
Db 541 GCTATCTCTGCTTAAGGGTTTGAAGTGGTGTAAAGGTGTCATTTGCTATCTCT 600
Qy 186 ProAlaValCySGInGluPheValAsnSerValPheCySGInGluGlnIleLeuValThr 205
Db 601 -----TACATCACT 609
Qy 206 GlnGluValValVal-----LeuArgGlyValAlaAsnPheAla 217

```

```
Db 610 GAGAACTAGATTAATGATGCTGCTATCTGATGCTGATTAACATTGCC 657
|||||
RESULT 10
US-09-206-942-44
; Sequence 44, Application US/09206942
; Patent No. 6432669
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: Protective Recombinant Haemophilus Influenzae High
; FILE REFERENCE: 1038-861 MIS:jb
; CURRENT APPLICATION NUMBER: US/09/206,942
; EARLIER FILING DATE: 1998-12-08
; EARLIER APPLICATION NUMBER: 09/167,568
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 44
; LENGTH: 3288
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
US-09-206-942-44

Alignment Scores:
Pred. No.: 1.67 Length: 3288
Score: 79.50 Matches: 48
Percent Similarity: 35.89% Conservative: 41
Best Local Similarity: 19.35% Mismatches: 90
Query Match: 7.10% Indels: 69
DB: Gaps: 9

US-10-034-500-2 (1-218) x US-09-206-942-44 (1-3288)
QY 11 ValGluSerTyrAsnTyrLeuValAspTyrSerGlySerMetMetLeuYshisValAla 30
Db 1063 ATGAAAGATTAACTTAATGCACTGGGGCAATATATCATTAACCAAGTTGCA 1122
QY 31 ValAArgGluProValIleGluLeuAlaIleGluValIleLeuYshisValAlaIleMet 50
Db 1123 GGTATTGATGTAATCTCA-----AAAGCCTTGTAGCAAT 1161
QY 51 ProLysMetSerTyrGlnGlyLeuTyrThrPheAla-----63
Db 1162 AAAAACAATACCTTTGAAGGGGCAATATCACCTTGACCGCATTAACCAATAGAA 1221
QY 64 ---ProTyrSerValIleIleProGlnGlySerTyrPheAsnSerCysValAlaGluCysAla 82
Db 1222 ATCAAAAGTAATATTAATCTGTTAAAGAGGCC---AATGTCAACCTTGAGCGCAAT 1278
QY 83 ValAsnThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGlyAspGly 102
Db 1279 TATGTAATGACAAATACACTTTAAGTAAGAAATCTCTAATAAAGCAAT---1335
QY 103 IleLysMetHisGluThrValIleAsnGlnMetProProGlnAlaIleLeuLeu 122
Db 1336 -----CTCACCGTT 1344
QY 123 ThrAspGlyHisAsnAsnLeuGlyMetAsnProValGluGluValLysSerIleTyrGln 142
Db 1345 ACCGGCTCGGCTATCAATATAGAAAAAATCTTAACGTTAAGTAAGTGTGTAAGTTTAA 1404
QY 143 ThrAsnProAsnValCysPheHisValIleSerPheAlaAspAsp-----157
Db 1405 GCTTAATCAAAATTAACGCTTTAAGTAATCGGCTATTGACAAACCAAGGCAAGTCAAC 1464
QY 158 -----AlaGluGlyLysAlaIleIleAspGlnIle-----167
Db 1465 ATTTCCATTGCCAAAGAGGGGCTCACTTTAAAGACATTATATACATAAGAGTTTAAAC 1524
QY 168 ValAlaIleAsnSerGlySer-----ValLeuValAspGly-----179
```

```
Db 1525 ATTAACCACTCGCACTCGCTTAACCGCATTAATTAAGGAACCAATATACCAACT 1584
|||
QY 180 -----LeuGlnLeuGlnAsnPro-----Ala 187
Db 1585 AACGGGATTTAATATATCATGATTAATAAATAACGCTGAATCAATGGGGCAAT 1644
QY 188 ValCysGlnGluProValAsnSerValPheCysGlnGluGlnIleLeuValThrGluGlu 207
Db 1645 ATCTGCAAAAGAAAGTAATCTCAAGATTCTTCGATTAATAATATCATTAACGACG 1704
QY 208 ValValValLeuAArgGlyValAsn 215
Db 1705 ATAAACAATCAAGAGGGTGAAT 1728

RESULT 11
US-09-206-942-42
; Sequence 42, Application US/09206942
; Patent No. 6432669
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: Protective Recombinant Haemophilus Influenzae High
; FILE REFERENCE: 1038-861 MIS:jb
; CURRENT APPLICATION NUMBER: US/09/206,942
; EARLIER FILING DATE: 1998-12-08
; EARLIER APPLICATION NUMBER: 09/167,568
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 42
; LENGTH: 3306
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
US-09-206-942-42

Alignment Scores:
Pred. No.: 1.69 Length: 3306
Score: 79.50 Matches: 48
Percent Similarity: 35.89% Conservative: 41
Best Local Similarity: 19.35% Mismatches: 90
Query Match: 7.10% Indels: 69
DB: Gaps: 9

US-10-034-500-2 (1-218) x US-09-206-942-42 (1-3306)
QY 11 ValGluSerTyrAsnTyrLeuValAspTyrSerGlySerMetMetLeuYshisValAla 30
Db 1081 ATGAAAGATTAACTTAATGCACTGGGGCAATATATCATTAACCAAGTTGCA 1140
QY 31 ValAArgGluProValIleGluLeuAlaIleGluValIleLeuYshisValAlaIleMet 50
Db 1141 GGTATTGATGTAATCTCA-----AAAGCCTTGTAGCAAT 1179
QY 51 ProLysMetSerTyrGlnGlyLeuTyrThrPheAla-----63
Db 1180 AAAAACAATACCTTTGAAGGGGCAATATCACCTTGACCGCATTAACCAATAGAA 1239
QY 64 ---ProTyrSerValIleIleProGlnGlySerTyrPheAsnSerCysValAlaGluCysAla 82
Db 1240 ATCAAAAGTAATATTAATCTGTTAAAGAGGCC---AATGTCAACCTTGAGCGCAAT 1296
QY 83 ValAsnThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGlyAspGly 102
Db 1297 TATGTAATGACAAATACACTTTAAGTAAGAAATGTAATTAAGGCAAT---1353
QY 103 IleLysMetHisGluThrValIleAsnGlnMetProProGlnAlaIleValIleLeuLeu 122
Db 1354 -----CTCACCGTT 1362
QY 123 ThrAspGlyHisAsnAsnLeuGlyMetAsnProValGluGluValLysSerIleTyrGln 142
```

```

Db 1363 ACCGGCTCCGCTATCAATATAGAAAAAATCTTACCGTGAAGGTAGTCTTAAGTTTAA 1422
Qy 143 ThrAenProAenValCyPheHisValSerPheAlaAenASP----- 157
Db 1423 GCTATTCGAAATTCAGCTTTACGATCCGCCCTATTGGACACCAAGGCAATGCAAC 1482
Qy 158 -----AlaGluGlyLysAlaIleIleAenGlnIle----- 167
Db 1483 ATTCGATTCGCAAGAGGGGCTCACTTAAAGCATTAAATACATTAAGATTAAAC 1542
Qy 168 ValAlaLeuAenSerGlySer-----ValLeuValAenGly----- 179
Db 1543 ATTACTACCAACTCCGCTCCGCTTACCGCACTATTATGAAGCAATATACCAACAGT 1602
Qy 180 -----LeuGlnLeuAenGlnAenPro-----Ala 187
Db 1603 AACGGGGATTAAATATATCATGATATATAAATAACGCTGAATAATCGCGCAAT 1662
Qy 188 ValCyGingluPheValAenSerValPheCyGingluGlnIleLeuValThrGluGlu 207
Db 1663 ATCTCGCAAAAAGAAAGATATCTCATGATTTCTCCGATTAATAATATATCATTAACAG 1722
Qy 208 ValValValLeuAenGlyValAen 215
Db 1723 ATAAATCAATCAAGAGGTGTTAAT 1746

```

## RESULT 12

```

US-08-470-179-89
Sequence 89, Application US/08470179
Patent No. 5645994
GENERAL INFORMATION:
APPLICANT: Huang Ph.D, Wei Mun
TITLE OF INVENTION: Method and Compositions for
NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESS: Trask, Britc and Rossa
STREET: P.O. Box 2550
CITY: Salt Lake City
STATE: Utah
COUNTRY: USA
ZIP: 84110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,179
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sweigert Ph.D, Susan E.
REGISTRATION NUMBER: 36,289
REFERENCE/DOCKET NUMBER: 2601
TELECOMMUNICATION INFORMATION:
TELEPHONE: 801-532-1922
TELEFAX: 801-531-9168
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 423 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: not relevant
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "gyra gene segment"
HYPOTHEICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Salmonella typhimurium
US-08-470-179-89

```

```

Alignment Scores:
Pred. No.: 0.0696 Length: 423
Score: 79.00 Matches: 25
Percent Similarity: 40.48% Conservative: 26
Best Local Similarity: 19.84% Mismatches: 59
Query Match: 7.06% Indels: 16
DB: 1 Gaps: 4

```

US-10-034-500-2 (1-218) x US-08-470-179-89 (1-423)

```

Qy 65 TyservaIleIleProGlnGlySerThrAenSerCyValAlaGluCyAlaValAen 84
Db 34 TACCGCATGAAGATTGGGCAATGACTGAAACAAAGCTTAATAAATCTGCC----- 87
Qy 85 ThrIleYserAenPleuGluIlePheGlyArgLeuThrProValGlyAenGlyIleLys 104
Db 88 CGTGTCTGGTGAC-----GTATCGTAAATATACATCCCGACGGGATTC-----GCA 138
Qy 105 MethAenGluThrValIleAenGlnMecProProGlnAlaAlaValIleLeuLeuThrAen 124
Db 139 GTGATGACACCATCGTTGATGCGGACGATTCGTCGCTGACATGCTGGTGAT 198
Qy 125 GlyHisAenAenGly-----MecAenPro 133
Db 199 GTCAGGGTAACTTCGTTCTATTGACGCGACATCCCGCGGCAATCGTTATACGAG 258
Qy 134 ValGluGluValIleYserIleTyrglnThrAenProAenValCyPheHisValAenSer 153
Db 259 ATCCGCTCGGCAAAATGCCACGAACTGATGCCCATCTCGAAAAAGACGGTGAT 318
Qy 154 PheAlaAenAenAlaGluGlyLysAlaIleIleAenGlnIleValAlaLeuAenSerGly 173
Db 319 TTCGTGATACATGATGACGAGGAAATAATCCGAGCATGCGGACCAACCAATTCG 378
Qy 174 SerValLeuValAenGly 179
Db 379 AATCGTGTGTAACGT 396

```

## RESULT 13

```

US-07-872-644-42/C
Sequence 42, Application US/07872644
Patent No. 5389527
GENERAL INFORMATION:
APPLICANT: Beavo, Joseph A.
APPLICANT: Bentley, Kelley
APPLICANT: Charbonneau, Harry
APPLICANT: Sonnenburg, William K.
TITLE OF INVENTION: DNA Encoding Mammalian
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray &
ADDRESS: Bicknell
STREET: Two First National Plaza, 20 South Clark
STREET: Street
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/872,644
FILING DATE: 19920420
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/688,356
FILING DATE: 04-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: No. 5389527and, Greta E.

```

```

; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/30822
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEEX: 25-3856
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3789 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 181..3006
; US-07-872-644-42

Alignment Scores:
Pred. No.: 2.47 Length: 3789
Score: 79.00 Matches: 37
Percent Similarity: 36.00% Conservative: 17
Best Local Similarity: 24.67% Mismatches: 50
Query Match: 7.06% Indels: 46
DB: 1 Gaps: 8

US-10-034-500-2 (1-218) x US-07-872-644-42 (1-3789)
Qy 90 LeuGIuIlePheGlyAArgLeu-----ThrProValGlyAArgGly 102
Db 2229 CTCGAGGTGAGTGGTGGAGTCCGAGTCTTGTAGACAGAGTAGCAGAGTGGGAGACAGA 2170
Qy 103 IleIysMetHisGluThrValIleAsnGlnMetProGlnAlaIleValIleLeu 122
Db 2169 AAAGGCTGATCCAGTGTGTGGAGGGGAGATCCG-----GAGCCCTTCTT 2122
Qy 123 ThrAArgGlyHisAsnAsnLeuGlyMetAsnProValGluGluValIlySerIleTyGln 142
Db 2121 -----CAC-----CATCAA 2113
Qy 143 ThrAsnProAsnValCys-----PheHisValVal 152
Db 2112 ACAGAACCCGGGCGAGTGTGGGCGAGTCAATTTGTAGTATTGATGAATTGATGCTCG 2053
Qy 153 SerPheAlaAspAspAlaGluGlyLysAlaIleIleAspGlnIleVal-----AlaLeu 170
Db 2052 CAGCATGCTCAGAGTGGCCATGGAAGTGTCTCTGGGCGAGAGCCAGAGAGTATGT 1993
Qy 171 AsnSerGlySerValLeuValaIleAspGlyLeuGlnLeuGlnAsnProAlaValCysGln 190
Db 1992 GAAACTGGCAAAAGTTGGAGTGCATGGCAGCCACAGAGCTGGAT-----CCCGTCATGGAG 1939
Qy 191 GluPhe-----ValaAsnSerValPheCysGlnGluGlnIleLeuValThrGln 206
Db 1938 AAGTTTGGTACTCTGCTCATCAGAGACCTTCATGCTGATCATCATCATTCATTCGCAAG 1879
Qy 207 GluValValValLeuValArgGlyValaAsnPhe 216
Db 1878 GTGGCTGCGACTGCTGCGCTCATTCATCTTT 1849

RESULT 14
US-08-297-494-42/C
; Sequence 42, Application US/08297494
; Patent No. 5580771
; GENERAL INFORMATION:
; APPLICANT: Beavo, Joseph A.
; APPLICANT: Bentley, Kelley
; APPLICANT: Charbonneau, Harry
; APPLICANT: Sonnenburg, William K.
; TITLE OF INVENTION: DNA Encoding Mammalian
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
```

```

; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; STREET: Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/297,494
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/688,356
; FILING DATE: 04-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5580771and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/30822
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEEX: 25-3856
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3789 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 181..3006
; US-08-297-494-42

Alignment Scores:
Pred. No.: 2.47 Length: 3789
Score: 79.00 Matches: 37
Percent Similarity: 36.00% Conservative: 17
Best Local Similarity: 24.67% Mismatches: 50
Query Match: 7.06% Indels: 46
DB: 1 Gaps: 8

US-10-034-500-2 (1-218) x US-08-297-494-42 (1-3789)
Qy 90 LeuGIuIlePheGlyAArgLeu-----ThrProValGlyAArgGly 102
Db 2229 CTCGAGGTGAGTGGTGGAGTCCGAGTCTTGTAGACAGAGTAGCAGAGTGGGAGACAGA 2170
Qy 103 IleIysMetHisGluThrValIleAsnGlnMetProGlnAlaIleValIleLeu 122
Db 2169 AAAGGCTGATCCAGTGTGTGGAGGGGAGATCCG-----GAGCCCTTCTT 2122
Qy 123 ThrAArgGlyHisAsnAsnLeuGlyMetAsnProValGluGluValIlySerIleTyGln 142
Db 2121 -----CAC-----CATCAA 2113
Qy 143 ThrAsnProAsnValCys-----PheHisValVal 152
Db 2112 ACAGAACCCGGGCGAGTGTGGGCGAGTCAATTTGTAGTATTGATGAATTGATGCTCG 2053
Qy 153 SerPheAlaAspAspAlaGluGlyLysAlaIleIleAspGlnIleVal-----AlaLeu 170
Db 2052 CAGCATGCTCAGAGTGGCCATGGAAGTGTCTCTGGGCGAGAGCCAGAGAGTATGT 1993
Qy 171 AsnSerGlySerValLeuValaIleAspGlyLeuGlnLeuGlnAsnProAlaValCysGln 190
Db 1992 GAAACTGGCAAAAGTTGGAGTGCATGGCAGCCACAGAGCTGGAT-----CCCGTCATGGAG 1939
```

Qy 191 Gluphe-----ValasServAlphecysgIngluInlleuValThrcIu 206  
Db 1938 AAGTTGGTACTGCTGATCAGAGACCTTCATGTCATCATCATCTATTGGCAAG 1879  
Qy 207 GluValValValValValValValValValValValValValValValVal 216  
Db 1878 GTGGCTGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1849

RESULT 15  
US-08-297-510-42/c  
Sequence 42: Application US/08297510  
Patent No. 5602019  
GENERAL INFORMATION:  
APPLICANT: Beavo, Joseph A.  
APPLICANT: Bentley, Kelley  
APPLICANT: Charbonneau, Harry  
APPLICANT: Sonnenburg, William K.  
TITLE OF INVENTION: DNA Encoding Mammalian  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
ADDRESSEE: Bicknell  
STREET: Two First National Plaza, 20 South Clark  
STREET: Street  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/297,510  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/688,356  
FILING DATE: 04-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5602019and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 27866/30822  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 346-5750  
TELEFAX: (312) 984-9740  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 42:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3789 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 181..3006  
US-08-297-510-42

Alignment Scores:  
Pred. No.: 2.47 Length: 3789  
Score: 79.00 Matches: 37  
Percent Similarity: 36.00% Conservative: 17  
Best Local Similarity: 24.67% Mismatches: 50  
Query Match: 7.06% Indels: 46  
DB: 1 Gaps: 8  
US-10-034-500-2 (1-218) x US-08-297-510-42 (1-3789)  
Qy 90 LeugluIlephgIyArgIeu-----ThrProValGlyAspGly 102

Db 2229 CTCGAGGATGTTGGTGTGAGCTTCAGGTTCTTGTAGACAGAGTACAGAGAGTGGAGACAGA 2170  
Qy 103 IlelysmethIsgIutrrValIleasngImeProProGlnIalValIleleu 122  
Db 2169 AAGCGGTCATCCAGTTGT 2122  
Qy 123 ThrAspGlyHisAsnAsnLeuGlyMetAsnProValGluGluValylsSerIleTyrcIu 142  
Db 2121 -----CAC-----CATCA 2113  
Qy 143 ThrAsnProAsnValCys-----PheHisValVal 152  
Db 2112 ACAGAACCGGCGCGAGTGTGCGGAGTCAATTTGTAGTATTGATGTAATTCATGCTCTG 2053  
Qy 153 SerPheAlaAspAspAlaGluGlyValAlaIleIleAspGlnIleVal-----AlaIeu 170  
Db 2052 CAGCATGCTCAGATGGCCATGAGTGTCTCTGCGGCAAGAGCAGAGAGTGTATGT 1993  
Qy 171 AsnSerGlySerValIleuValAspGlyLeuGlnIleuGlnAsnProAlaValCysGln 190  
Db 1992 GAACCTGGCAAGTTGGAGTGTGATGCGACGACAGCGCTGAT-----CCGTCATGGAG 1939  
Qy 191 Gluphe-----ValasServAlphecysgIngluInlleuValThrcIu 206  
Db 1938 AAGTTGGTACTGCTGATCAGAGACCTTCATGTCATCATCATCTATTGGCAAG 1879  
Qy 207 GluValValValValValValValValValValValValValValValVal 216  
Db 1878 GTGGCTGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1849

RESULT 16  
US-08-479-532-42/c  
Sequence 42: Application US/08479532  
Patent No. 576752  
GENERAL INFORMATION:  
APPLICANT: Beavo, Joseph A.  
APPLICANT: Bentley, Kelley  
APPLICANT: Charbonneau, Harry  
APPLICANT: Sonnenburg, William K.  
TITLE OF INVENTION: DNA Encoding Mammalian  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
ADDRESSEE: Bicknell  
STREET: Two First National Plaza, 20 South Clark  
STREET: Street  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/479,532  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/297,494  
FILING DATE:  
APPLICATION NUMBER: US 07/688,356  
FILING DATE: 04-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 576752and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 27866/30822  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 346-5750  
TELEFAX: (312) 984-9740



```

      TELEEX    25-3856
      INFORMATION FOR SEQ ID NO: 42:
      SEQUENCE CHARACTERISTICS:
        LENGTH: 3789 base pairs
        TYPE: nucleic acid
        STRANDEDNESS: single
        TOPOLOGY: linear
        MOLECULE TYPE: cDNA
      FEATURE:
        NAME/KEY: CDS
        LOCATION: 181..3006
US-08-479-532-42
Alignment Scores:
Pred. No.:          2.47           Length:         3789
Score:              79.00          Matches:         37
Percent Similarity: 36.00%         Conservative:   17
Best Local Similarity: 24.67%       Mismatches:     50
Query Match:         7.06%          Indels:         46
                                      Gaps:            8
DB:                               1
US-10-034-500-2 (1-218) x US-08-479-532-42 (1-3789)
Qy      90 LeugluilephegiylArgLeu-----ThrProvalGIlyAspgIy 102
Db      2229 CTCGAGTCAGTGGTTGTGAAGCTCCACGATTCCTTGTAAGCAGGTAGCACAAATCGGAGAACA 2170
Qy      103 IlelysMechisglutlrivalIIleasnGlmetPrpProglnAlalavalIIleuleu 122
Db      2169 AAAGCGGCATCCAGTTGTGTGGTGGGGGATCCCC-----GTAGCCCCTTCT 2122
Qy      123 ThrAspGLYhsaaNanInleUglyMetAsnProValGluGluValysSerIleTryGln 142
Db      2121 -----CAC-----CATCA 2113
Qy      143 ThrasPProasValCys-----PheHIsValVal 152
Db      2112 ACAgAACCGGCCcAgTgcggGcAGtCaATTtttgTaTtgaTtaGaAAtTcaTgtctcg 2055
Qy      153 SerPheaLaaspAlagLugllyVlsalaIlleAspGlnlleval-----Alaleu 170
Db      2052 CAGCATGCTCAGATGgCcArGatGcAGAAgtGcatCtccGcAGAcGAGcAGAGtGATGT 1993
Qy      171 AsnsErGlyServalLeuValAspGlyLeuGlnleuLeuGlnInsProAlaValCysGln 190
Db      1992 GAACATGcCAAAgTTGAGtGcAGtcAGcCAGcCAcAGcGTGgAT-----CCcGTCArtGGAG 1935
Qy      191 GlupHe-----ValAsnsErValPhecysGlnGluGlnleuValThrGlu 206
Db      1938 AAGTTGtGTACTCtGcTcATATAGAGAcCTTCaTgStgTATcATcATcATcATtGcCAAG 1875
Qy      207 GluValValValLeuArgIlyValAsnPhe 216
Db      1878 GTGCTGCGATACTGCGCCCTCATTCACATT 1849
RESULT 17
US-08-455-526-42/c
; Sequence 42, Application US/0845526
; Patent No. 5789553
GENERAL INFORMATION:
APPLICANT: Beavo, Joseph A.
APPLICANT: Bentley, Kelley
APPLICANT: Charbonneau, Harry
APPLICANT: Sonnenburg, William K.
TITLE OF INVENTION: DNA Encoding Mammalian
TITLE OF INVENTION: Phosphodiesterases
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESSSEE: Bicknell
STREET: Two First National Plaza, 20 South Clark
CITY: Chicago
```

```

1 STATE: Illinois
2 COUNTRY: USA
3 ZIP: 60603
4
5 COMPUTER READABLE FORM:
6 MEDIUM TYPE: Floppy disk
7 COMPUTER: IBM PC compatible
8 OPERATING SYSTEM: PC-DOS/MS-DOS
9 SOFTWARE: Patentin Release #1.0, Version #1.25
10
11 CURRENT APPLICATION DATA:
12 APPLICATION NUMBER: US/08/455,526
13 FILING DATE: 31-MAY-1995.
14 CLASSIFICATION: 530
15
16 PRIOR APPLICATION DATA:
17 APPLICATION NUMBER: US 08/297,494
18 FILING DATE: 29-AUG-1994
19 APPLICATION NUMBER: US 07/668,356
20 FILING DATE: 04-APR-1991
21
22 ATTORNEY/AGENT INFORMATION:
23 NAME: No. 578955Iand, Greta E.
24 REGISTRATION NUMBER: 35,302
25 REFERENCE/DOCKET NUMBER: 27866/30822
26
27 TELECOMMUNICATION INFORMATION:
28 TELEPHONE: (312) 346-5750
29 TELEFAX: (312) 984-9740
30
31 TELEX: 25-3856
32
33 INFORMATION FOR SEO. ID NO: 42:
34
35 SEQUENCE CHARACTERISTICS:
36 LENGTH: 3789 base pairs
37 TYPE: nucleic acid
38 STRANDEDNESS: single
39 TOPOLOGY: linear
40
41 MOLECULE TYPE: cDNA
42
43 FEATURE:
44 NAME/KEY: CDS
45 LOCATION: 181..3006
46
47 US-08-455-526-42
48
49
50
51 Alignment Scores:
52 Pred. No.: 2.47 Length: 3789
53 Score: 79.00 Matches: 37
54 Percent Similarity: 36.00% Conservative: 17
55 Best Local Similarity: 24.67% Mismatches: 50
56 Query Match: 7.06% Indels: 46
57 Gaps: 8
58
59 US-10-034-500-2 (1-218) x US-08-455-526-42 (1-3789)
60
61 QY 90 LeuGluLeuPheGlyArgLeu-----ThiProValGlyAspGly 102
62 |||||:|:| ||| |||||
63 Db 2229 CTCAGAGTACTGTGTAGAGCTCCAGGCTTCTTGAGACAGCTGACAGAGTGGAGACAGCA 2170
64
65 QY 103 IleLysMetHisGluThrValIleLeaGlnMetProProGlnAlaAlaValIleLeuLeu 122
66 ::|||:|:| |||||:|:| |||||
67 Db 2169 AAAGCGCGCATCCAGTGTGTGTAGGAGGGGAGATCCG-----GTAGCCCTTCTT 2122
68
69 QY 123 ThrAspGlyHisAsnAsnLeuGlyMetAsnProValGluGluValLysSerIleTyrGln 142
70 |||
71 Db 2121 -----CAC-----CATCA 2113
72
73 QY 143 ThrAsnProAsnValCys-----PheHisValVal 152
74 |||||:|:| ||| |||||:|:|
75 Db 2112 ACAAGACCGGGCCAGTGTCCGGCGCAGTCAATTTTGTAGTATTGATGAAATTCATCTCCG 2055
76
77 QY 153 SerPheAlaAspAlaGluGlyLysAlaIleIleAspGlnIleVal-----AlaLeu 170
78 |||||:|:| |||||:|:| |||||
79 Db 2052 CAGATGCTCAGATGGCCATCGAAGTGCATCTCTCGGCGAGAGCGAGGTATGT 1993
80
81 QY 171 AsnSerGlySerValLeuValAspGlyLeuGlnLeuLeuGlnAsnProAlaValCysGln 190
82 ::|||:|:| |||||:|:| |||||
83 Db 1992 GAAATCGGCAAAATTGGAGTGCATGGCAGCCACAGGCTGGAT-----CCCTCATGGAG 1939
84
85 QY 191 GluPhe-----ValAsnSerValPheCysGlnGluGlnIleLeuValThrGlu 206
86 ::|||:|:| ::|||:|:|

```

```
Db      1938 AAGTTGTGACTCTGCATCACAACCTTCATCGTGATCATCATCTCAATTGGCAAG   187
Oy      207 GUUVA|VAlValleatrggYvAlasmphe 216
          ::|||::|:|||||
Db      1878 GTGGCTGCCAATACCTCCATCACCTT    1849

RESULT 18
US-08-455-525-42/C
; Sequence 42, Application US/08455525
; Patent No. 5800987
; GENERAL INFORMATION:
; APPLICANT: Beavo, Joseph A.
; APPLICANT: Bentley, Kelley
; APPLICANT: Charbonneau, Harry
; APPLICANT: Somendure, William K.
; TITLE OF INVENTION: DNA Encoding Mammalian
; NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; STREET: Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
ZIP: 60603

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/455,525
FILING DATE: 31-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/297,494
FILING DATE:
APPLICATION NUMBER: US 07/688,356
FILING DATE: 04-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: No. 5800987and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/30822
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 346-5750
TELEFAX: (312) 984-9740
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 3789 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
; NAME/KEY: CDS
; LOCATION: 181..3006
US-08-455-525-42

Alignment Scores:
Pred. No.:                2.47           Length:       3789
Score:                    79.00           Matches:        37
Percent Similarity:      36.00%           Conservative:   17
Best Local Similarity:   24.67%           Mismatches:    50
Query Match:             7.06%            Indels:         46
DB:                      1               Gaps:           8

US-10-034-500-2 (1-218) x US-08-455-525-42 (1-3789)
Oy      90 leuguuiehpqgyaglu-----ThrProValgllyapgly 102
          |||||::|:|||
```

```

Db 2229 CTGAGGTAAGTGGTAGACCTCCAGGCTTCTTTGTAAGACAGAGTAGACAGAAAGTGGAGACAGA 2170
QY 103 IleywMechisgluThrvall1leaenGlnMetProFroGlnAlaVal1lleuLeu 122
Db 2169 AAAGCGCTGACATCCAGTTGGTAGAGGGGAGATCCCG-----GTAGCCCTTCTT 2122
QY 123 ThrspglYlHisanhenLeuGlyMetanProValGlUGluValLyserr1eYrGln 142
Db 2121 -----CAC-----CATCAA 2113
QY 143 ThrAnProAnValCys-----PheHisValVal 152
Db 2112 ACAAGACCGGCCAGGTCTCCGGACAGTCATTTTGTAGTTATGTAGTAAGAAATTCATCTCTG 2053
QY 153 SerPheAlaaspAapalagIugIyVala1le1leaspGln1leVal-----AlaLeu 170
Db 2052 CAGCATCTCAGAGTAGGCGCATGGAAGTGCATCTCTCCGGCAGAGACGAGAGTGTATGT 1993
QY 171 AsnserGlyserValleuValaapGlyLeuGlnLeuGlnAsnProAlaValCysGln 190
Db 1992 GAAACTGGCAAAAGTTGGAGTCGATGCGACGACCAAGGCTGGAT-----CCCGTCATGGAG 1939
QY 191 GluPhe-----ValAsnserValPheCysGlnGluGln1leValThrnGlu 206
Db 1938 AAGTTTGCTGATCTCGTCATCAGACAGACCTTCATGTGTGATCATCATCATCTCTTGGCAG 1879
QY 207 GluValValleuAArgIyValaenPhe 216
Db 1878 GTGGCTGCGATCACTGGCGCCCTCATTCCTTT 1849

RESULT 19
US-09-139-491-42/c
; Sequence 42, Application US/09139491
; Patent No. 6015677
; GENERAL INFORMATION:
; APPLICANT: Beavo, Joseph A.
; APPLICANT: Bentley, Kelley
; APPLICANT: Charbonneau, Harry
; APPLICANT: Sonnenburg, William K.
; TITLE OF INVENTION: DNA Encoding Mammalian
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; STREET: Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Releasee #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/139,491
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,525
; FILING DATE: 31-MAY-1995
; APPLICATION NUMBER: 08/297,494
; FILING DATE:
; APPLICATION NUMBER: US 07/668,356
; FILING DATE: 04-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6015677and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 2766/30822
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750

```

TELEFAX: (312) 984-9740  
 TELEX: 25-3856  
 INFORMATION FOR SEQ ID NO: 42:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3789 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 181..3006  
 US-09-139-491-42

Alignment Scores:  
 Pred. No.: 2.47 Length: 3789  
 Score: 79.00 Matches: 37  
 Percent Similarity: 36.00% Conservative: 17  
 Best Local Similarity: 24.67% Mismatches: 50  
 Query Match: 7.06% Indels: 46  
 DB: 3 Gaps: 8

US-10-034-500-2 (1-218) x US-09-139-491-42 (1-3789)

```

QY 90 LeuGluIlePheGlyArgLeu-----ThrProValGlyAspGly 102
DB 2229 CTCGAGGTAGTGTGAGCTCCAGGTTCTTGAGACAGGTAGCAGAGTGGAGACAGA 2170
QY 103 IleLysMetHisGluThrValIleAsnGlnMetProProGlnAlaIleValIleLeu 122
DB 2169 AAGGGGTGATCCAGCTTGAGTGGAGGGGATCCG-----GTAGCCCTTCTT 2122
QY 123 ThrAspGlyHisAsnAsnLeuGlyMetAsnProValGluGluValLysSerIleTyrGln 142
DB 2121 -----CAC-----CATCAA 2113
QY 143 ThrAsnProAsnValCys-----PheHisValVal 152
DB 2112 ACAGAACCCGGCCAGTGGCGGCGACATCAATTTGTATTGATGAGAAATTCATGTCCTG 2053
QY 153 SerPheAlaAspAspAlaGluGlyLysAlaIleIleAspGlnIleVal-----AlaLeu 170
DB 2052 CAGCATGCTCAGAGTGGCCAGTGAAGTGCATCTCGCGGAGGAGCGAGAGCTGATGT 1993
QY 171 AsnSerGlySerValLeuValAspGlyLeuGlnLeuGlnAsnProAlaValCysGln 190
DB 1992 GAACTGGCGAAAGTTGAGATGCATGCGCAGCCAGCGCTGAT-----CCGTCATGGAG 1939
QY 191 GluPhe-----ValAsnSerValPheCysGlnGluGlnIleLeuValThrGln 206
DB 1938 AGTTGGTGTACTGTCGTCATCAGACCTTCATGTTGATCATCATCATCTCATTTGGCAAG 1879
QY 207 GluValValValLeuArgGlyValAsnPhe 216
DB 1878 GTGGCTGCATACTGCGCCCTCATCTT 1849

```

RESULT 20

PCT-US92-03222-42/c  
 Sequence 42, Application PC/TUS92032222  
 GENERAL INFORMATION:

APPLICANT: Beavo, Joseph A.  
 APPLICANT: Bentley, Kelley  
 APPLICANT: Charbonneau, Harry  
 APPLICANT: Sonnenburg, William K.  
 TITLE OF INVENTION: DNA Encoding Mammalian  
 TITLE OF INVENTION: Phosphodiesterases  
 NUMBER OF SEQUENCES: 58  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
 ADDRESSEE: Bicknell  
 STREET: Two First National Plaza, 20 South Clark  
 STREET: Street  
 CITY: Chicago

STATE: Illinois  
 COUNTRY: USA  
 ZIP: 60603  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US92/03222  
 FILING DATE: 19920420  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/688,356  
 FILING DATE: 04-APR-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Noland, Greta E.  
 REGISTRATION NUMBER: 35,302  
 REFERENCE/DOCKET NUMBER: 27866/30822  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (312) 346-5750  
 TELEFAX: (312) 984-9740  
 TELEX: 25-3856  
 INFORMATION FOR SEQ ID NO: 42:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3789 base pairs  
 TYPE: NUCLEIC ACID  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 181..3006  
 PCT-US92-03222-42

Alignment Scores:  
 Pred. No.: 2.47 Length: 3789  
 Score: 79.00 Matches: 37  
 Percent Similarity: 36.00% Conservative: 17  
 Best Local Similarity: 24.67% Mismatches: 50  
 Query Match: 7.06% Indels: 46  
 DB: 5 Gaps: 8

US-10-034-500-2 (1-218) x PCT-US92-03222-42 (1-3789)

```

QY 90 LeuGluIlePheGlyArgLeu-----ThrProValGlyAspGly 102
DB 2229 CTCGAGGTAGTGTGAGCTCCAGGTTCTTGAGACAGGTAGCAGAGTGGAGACAGA 2170
QY 103 IleLysMetHisGluThrValIleAsnGlnMetProProGlnAlaIleValIleLeu 122
DB 2169 AAGGGGTGATCCAGCTTGAGTGGAGGGGATCCG-----GTAGCCCTTCTT 2122
QY 123 ThrAspGlyHisAsnAsnLeuGlyMetAsnProValGluGluValLysSerIleTyrGln 142
DB 2121 -----CAC-----CATCAA 2113
QY 143 ThrAsnProAsnValCys-----PheHisValVal 152
DB 2112 ACAGAACCCGGCCAGTGGCGGCGACATCAATTTGTATTGATGAGAAATTCATGTCCTG 2053
QY 153 SerPheAlaAspAspAlaGluGlyLysAlaIleIleAspGlnIleVal-----AlaLeu 170
DB 2052 CAGCATGCTCAGAGTGGCCAGTGAAGTGCATCTCGCGGAGGAGCGAGAGCTGATGT 1993
QY 171 AsnSerGlySerValLeuValAspGlyLeuGlnLeuGlnAsnProAlaValCysGln 190
DB 1992 GAACTGGCGAAAGTTGAGATGCATGCGCAGCCAGCGCTGAT-----CCGTCATGGAG 1939
QY 191 GluPhe-----ValAsnSerValPheCysGlnGluGlnIleLeuValThrGln 206
DB 1938 AGTTGGTGTACTGTCGTCATCAGACCTTCATGTTGATCATCATCATCTCATTTGGCAAG 1879

```

Qy 207 GIUVAIValleuArglyValAsnPhe 216  
Db 1878 GTGGCTGCATACGTGGCTCATTCACCTT 1849

RESULT 21  
US-07-872-644-38/c  
Sequence 38, Application US/07872644  
Patent No. 5389527  
GENERAL INFORMATION:  
APPLICANT: Beavo, Joseph A.  
APPLICANT: Bentley, Kelley  
APPLICANT: Charbonneau, Harry  
APPLICANT: Sonnenburg, William K.  
TITLE OF INVENTION: DNA Encoding Mammalian  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
ADDRESSEE: Bicknell  
STREET: Two First National Plaza, 20 South Clark  
STREET: Street  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/872,644  
FILING DATE: 19920420  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/688,356  
FILING DATE: 04-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5389527and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 27866/30822  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 346-5750  
TELEFAX: (312) 984-9740  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 38:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4131 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 148..2910  
US-07-872-644-38

Alignment Scores:  
Pred. No.: 2.85 Length: 4131  
Score: 79.00 Matches: 37  
Percent Similarity: 36.00% Conservative: 17  
Best Local Similarity: 24.67% Mismatches: 50  
Query Match: 7.06% Indels: 46  
DB: 1 Gaps: 8  
US-10-034-500-2 (1-218) x US-07-872-644-38 (1-4131)

Qy 90 leugliuepneGlyArgleu-----ThProValGlyAspGly 102  
Db 2133 CTCGAGTGAAGTTGCTGACGCTTCCTTGTAGACAGCTGACGAAGTGGAGACAGCA 2074  
Qy 103 TleuysMetHleGluThrValIleAsnGlnMetProProGlnAlaAlaValIleleu 122

Db 2073 AAAGCGTCATCCAGTTGTGTAGAGGGGATCCCG-----GTAGCCCTTCTT 2026  
Qy 123 ThrAspGlyHisAsnAsnLeuGlyMetAsnProValGluGluValIysSerIleTyrGln 142  
Db 2025 -----CAC-----CATCA 2017  
Qy 143 ThrAsnProAsnValCys-----PheHisValVal 152  
Db 2016 ACAGAACCGGCGCAGTGTCCGGCAGTCAATTTGTGATTTATGTGAATAATCATGCTCTG 1957  
Qy 153 SerPheAlaAspAspAlaGluGlyValAlaIleIleAspGlnIleVal-----AlaLeu 170  
Db 1956 CAGCATGCTCAGATGGCCATGAGTCTATCTCCGGCAGAGACGAGAGTGTATGT 1897  
Qy 171 AsnSerGlySerValIleuValAspGlyLeuGlnIleuGlnAsnProAlaValCysGln 190  
Db 1896 GAACCTGCAGAAAGTTGAGTGCATGAGCCAGCAGGCTGAT-----CCGTCATGGAG 1843  
Qy 191 GluPhe-----ValAsnSerValPheCysGlnGlnGlnIleValThrGln 206  
Db 1842 AAGTTGGTGTACTGTCATCAGAGACCTTCATGTGTATCATCATCTCATTCAGCAAG 1783  
Qy 207 GIUVAIValleuArglyValAsnPhe 216  
Db 1782 GTGGCTGCATACGTGGCTCATTCACCTT 1753

RESULT 22  
US-08-297-494-38/c  
Sequence 38, Application US/08297494  
Patent No. 5580771  
GENERAL INFORMATION:  
APPLICANT: Beavo, Joseph A.  
APPLICANT: Bentley, Kelley  
APPLICANT: Charbonneau, Harry  
APPLICANT: Sonnenburg, William K.  
TITLE OF INVENTION: DNA Encoding Mammalian  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
ADDRESSEE: Bicknell  
STREET: Two First National Plaza, 20 South Clark  
STREET: Street  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/297,494  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/688,356  
FILING DATE: 04-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5580771and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 27866/30822  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 346-5750  
TELEFAX: (312) 984-9740  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 38:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4131 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear



Patent No. 5776752  
GENERAL INFORMATION:  
APPLICANT: Beavo, Joseph A.  
APPLICANT: Bentley, Kelley  
APPLICANT: Charbonneau, Harry  
APPLICANT: Sonnenburg, William K.  
TITLE OF INVENTION: DNA Encoding Mammalian  
TITLE OF INVENTION: Phosphodiesterases  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
ADDRESSEE: Bicknell  
STREET: Two First National Plaza, 20 South Clark  
STREET: Street  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/479,532  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/297,494  
FILING DATE:  
APPLICATION NUMBER: US 07/688,356  
FILING DATE: 04-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5776752and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 27866/30822  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 346-5750  
TELEFAX: (312) 984-9740  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 38:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4131 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 148..2910  
US-08-479-532-38  
Alignment Scores:  
Pred. No.: 2.85 Length: 4131  
Score: 79.00 Matches: 37  
Percent Similarity: 36.00% Conservative: 17  
Best Local Similarity: 24.67% Mismatches: 50  
Query Match: 7.06% Indels: 46  
DB: 1 Gaps: 8  
US-10-034-500-2 (1-218) x US-08-479-532-38 (1-4131)  
Qy 90 LeugluIlePhegLYAArgLeu-----Th-ProValGlyAspGly 102  
Db 2133 CTCAGAGTAGTTGAGCTCCAGGTTCTTGTAGACAGAGTTCAGAAAGTGGAGACAGA 2074  
Qy 103 IleYsMeChIsgLurThValIleAnGImetProProGlnAlaIleValIleleu 122  
Db 2073 AAAGCGTCATCCAGTTGTGTAGAGGGGAGATCCCG-----GTAGCCCTTCTT 2026  
Qy 123 ThrAepGlyHleAsnAsnLeuGlyMetAenProValGluGluValIlyserIleYrGln 142  
Db 2025 -----CAC-----CATCAA 2017

Qy 143 ThrAenProAenValCyG-----PheHleValVal 152  
Db 2016 ACAGAACCGGCGCAGTGTCCGGCAGTCAATTTTGAATTATGTGAATTCATGTCTCTG 1957  
Qy 153 SerPheAlaAspAapAlaGluGlyLyValAlaIleIleAepGlnIleVal-----AlaLeu 170  
Db 1956 CAGCATGTCCAGATGGCAGTGCAGATGTCATCTCCGGCAGAGAGCGAGAGTGTATGT 1897  
Qy 171 AenSerGlySerValIleValAspGlyLeuGlnLeuGlnAenProAlaValCyGgln 190  
Db 1896 GAACTGGCAAGTTGAGTGCATGAGTGCAGCCAGCGCTGAT-----CCGTCATGGAG 1843  
Qy 191 GlupHe-----ValAenSerValPheCyGglnGlnIleLeuValIthrGln 206  
Db 1842 AAGTTTGTGTACTGTCTCATCAGACCTTCATGTGTATCATCATCATTCATTCGCAAG 1783  
Qy 207 GluValValValLeuArgGlyValAsnPhe 216  
Db 1782 GTGCTGCGATATCGCGCTCATTCACATT 1753  
RESULT 25  
US-08-455-526-38/c  
Sequence 38, Application US/08455526  
GENERAL INFORMATION:  
APPLICANT: Beavo, Joseph A.  
APPLICANT: Bentley, Kelley  
APPLICANT: Charbonneau, Harry  
APPLICANT: Sonnenburg, William K.  
TITLE OF INVENTION: DNA Encoding Mammalian  
TITLE OF INVENTION: Phosphodiesterases  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
ADDRESSEE: Bicknell  
STREET: Two First National Plaza, 20 South Clark  
STREET: Street  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/455,526  
FILING DATE: 31-MAY-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/297,494  
FILING DATE: 29-AUG-1994  
APPLICATION NUMBER: US 07/688,356  
FILING DATE: 04-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5789553and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 27866/30822  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 346-5750  
TELEFAX: (312) 984-9740  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 38:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4131 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS

```

LOCATION: 148..2910
US-08-455-526-38

Alignment Scores:
Pred. No.: 2.85      Length: 4131
Score: 79.00      Matches: 37
Percent Similarity: 36.00%      Conservative: 17
Best Local Similarity: 24.67%      Mismatches: 50
Query Match: 7.06%      Indels: 46
DB: 1      Gaps: 8

US-10-034-500-2 (1-218) x US-08-455-526-38 (1-4131)

Qy 90 LeuGlulIephGlyArgLeu-----ThrProValGlyAspGly 102
Db 2133 CTCGAGGTAGTGGTGTGAGCTCCAGGTTCTTGTAGACAGAGTAGACAGAGTGGAGACAGCA 2074
Qy 103 IleYsMetHisGluThrValIleAsnGlnMetProProGlnAlaIleValIleLeuLeu 122
Db 2073 AAAGCGTCATCCAGATTGTGTGAGGGGAGATCCCG-----GTAGCCCTTCTT 2026
Qy 123 ThrAspGlyHisAsnAsnLeuGlyMetAsnProValGluGluValIlySerIleTyrgln 142
Db 2025 -----CAC-----CATCAA 2017
Qy 143 ThrAsnProAsnValCys-----PheHisValVal 152
Db 2016 ACAGAAACCGGCGCAGGTGTGGGCGACATCAATTTGTAGTATTGATGAATTCATGTCCTG 1957
Qy 153 SerPheAlaAspAspAlaGluGlyLysAlaIleIleAspGlnIleVal-----AlaLeu 170
Db 1956 CAGCATGTCTCAGAGATGGCCATGGAAGTGCATCTCCGCGCAGAGACGAGAGTGTATGT 1897
Qy 171 AsnSerGlySerValLeuValAspGlyLeuGlnLeuGluGlnAsnProAlaValCysGln 190
Db 1896 GAAACTGGCAAAATTGAGATGCATGCGACCCACAGGCTGGAT-----CCCGTCATGGAG 1843
Qy 191 GluPhe-----ValAsnSerValPheCysGlnGluGlnIleLeuValThrgln 206
Db 1842 AGCTTGGTACTCTGCTCATCAGACCTTCATGCTGATCATCATCATCTCATGTGGCAAG 1783
Qy 207 GluValValValLeuArgGlyValAsnPhe 216
Db 1782 GTGGCTGCATACTGCGCCTCATTTACTTT 1753

RESULT 26
US-08-455-525-38/c
; Sequence 38, Application US/08455525
; GENERAL INFORMATION:
; APPLICANT: Beavo, Joseph A.
; APPLICANT: Bentley, Kelley
; APPLICANT: Charbonneau, Harry
; APPLICANT: Sonnenburg, William K.
; TITLE OF INVENTION: DNA Encoding Mammalian
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,525

```

```

FILING DATE: 31-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/297,494
FILING DATE:
APPLICATION NUMBER: US 07/688,356
FILING DATE: 04-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: No. 5800987and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/30822
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 346-5750
TELEFAX: (312) 984-9740
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 4131 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURES:
NAME/KEY: CDS
LOCATION: 148..2910
US-08-455-525-38

Alignment Scores:
Pred. No.: 2.85      Length: 4131
Score: 79.00      Matches: 37
Percent Similarity: 36.00%      Conservative: 17
Best Local Similarity: 24.67%      Mismatches: 50
Query Match: 7.06%      Indels: 46
DB: 1      Gaps: 8

US-10-034-500-2 (1-218) x US-08-455-525-38 (1-4131)

Qy 90 LeuGlulIephGlyArgLeu-----ThrProValGlyAspGly 102
Db 2133 CTCGAGGTAGTGGTGTGAGCTCCAGGTTCTTGTAGACAGAGTAGACAGAGTGGAGACAGCA 2074
Qy 103 IleYsMetHisGluThrValIleAsnGlnMetProProGlnAlaIleValIleLeuLeu 122
Db 2073 AAAGCGTCATCCAGATTGTGTGAGGGGAGATCCCG-----GTAGCCCTTCTT 2026
Qy 123 ThrAspGlyHisAsnAsnLeuGlyMetAsnProValGluGluValIlySerIleTyrgln 142
Db 2025 -----CAC-----CATCAA 2017
Qy 143 ThrAsnProAsnValCys-----PheHisValVal 152
Db 2016 ACAGAAACCGGCGCAGGTGTGGGCGACATCAATTTGTAGTATTGATGAATTCATGTCCTG 1957
Qy 153 SerPheAlaAspAspAlaGluGlyLysAlaIleIleAspGlnIleVal-----AlaLeu 170
Db 1956 CAGCATGTCTCAGAGATGGCCATGGAAGTGCATCTCCGCGCAGAGACGAGAGTGTATGT 1897
Qy 171 AsnSerGlySerValLeuValAspGlyLeuGlnLeuGluGlnAsnProAlaValCysGln 190
Db 1896 GAAACTGGCAAAATTGAGATGCATGCGACCCACAGGCTGGAT-----CCCGTCATGGAG 1843
Qy 191 GluPhe-----ValAsnSerValPheCysGlnGluGlnIleLeuValThrgln 206
Db 1842 AGCTTGGTACTCTGCTCATCAGACCTTCATGCTGATCATCATCATCTCATGTGGCAAG 1783
Qy 207 GluValValValLeuArgGlyValAsnPhe 216
Db 1782 GTGGCTGCATACTGCGCCTCATTTACTTT 1753

RESULT 27
US-09-139-491-38/c
; Sequence 38, Application US/09139491
; Patent No. 6015677

```

GENERAL INFORMATION:  
APPLICANT: Beavo, Joseph A.  
APPLICANT: Bentley, Kelley  
APPLICANT: Charbonneau, Harry  
APPLICANT: Sonnenburg, William K.  
TITLE OF INVENTION: DNA Encoding Mammalian  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
STREET: Two First National Plaza, 20 South Clark  
STREET: Street  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/139,491  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/455,525  
FILING DATE: 31-MAY-1995  
APPLICATION NUMBER: 08/297,494  
FILING DATE:  
APPLICATION NUMBER: US 07/688,356  
FILING DATE: 04-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 6015672and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 27866/30822  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 346-5750  
TELEFAX: (312) 984-9740  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 38:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4131 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 148..2910  
US-09-139-491-38  
Alignment Scores:  
Pred. No.: 2.85 Length: 4131  
Score: 79.00 Matches: 37  
Percent Similarity: 36.00% Conservative: 17  
Best Local Similarity: 24.67% Mismatches: 50  
Query Match: 7.05% Indels: 46  
DB: 3 Gaps: 8  
US-10-034-500-2 (1-218) x US-09-139-491-38 (1-4131)  
Qy 90 LeugluIlePhagIyAryleu-----ThrProValGlyApsgly 102  
Db 2133 CTGAGGAGTAGTTGGAGACTTCAGGTTCTTGTAGACAGGTAGACAGAAAGGGAGACAGCA 2074  
Qy 103 IleYsMeTrIsgIuThrValIleAenGInMeTProGInAlaIaValIleuLeu 122  
Db 2073 AAAGCGGTGATCCAGTGTGTGGTAGGGGAGATCCCG-----GTAGCCCTTCTT 2026  
Qy 123 ThrAepGlyHIsAenAenLeuGlyMeTAsnProValGIuGIuValIySerIeTyrcIn 142

Db 2025 -----CAC-----CATCA 2017  
Qy 143 ThrAsnProAenValCyS-----PheHisValVal 152  
Db 2016 ACAGAACCGGGCCAGTGTGCGGCGAGTCATATTGTGATTGATGAATTCATGTCTCG 1957  
Qy 153 SerPheAlaAspAapAlaGlyIyValAlleIleIleAspGInIleVal-----AlaLeu 170  
Db 1956 CAGCATGCTCAGATGCGCAGTAGGAAGTGTCTCTCGGCGCAGAGAGGAGGTGTATGT 1897  
Qy 171 AsnSerGlySerValIleuValApsGlyIleuGInIleuGInAenProAlaValCySgIn 190  
Db 1896 GAACGTGCAGAAAGTTGAGATGATGCGACCCACAGCGTGAAT-----CCCGTATGGAG 1843  
Qy 191 Gluphe-----ValAsnSerValPheCySgInGIuGIuIleuValThrGIu 206  
Db 1842 AAGTTGTGTACTCGTCATAGAGACCTTATGTGTGATCATCATCTATTTGGCAAG 1783  
Qy 207 GIuValValValIleuAArgGlyValAsnPhe 216  
Db 1782 GTGGCTGCGATACGTGCGCTCATTCACCTT 1753  
RESULT 28  
PCT-US92-03222-38/c  
Sequence 38, Application PC/TUS9203222  
GENERAL INFORMATION:  
APPLICANT: Beavo, Joseph A.  
APPLICANT: Bentley, Kelley  
APPLICANT: Charbonneau, Harry  
APPLICANT: Sonnenburg, William K.  
TITLE OF INVENTION: DNA Encoding Mammalian  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
STREET: Two First National Plaza, 20 South Clark  
STREET: Street  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/03222  
FILING DATE: 19920420  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/688,356  
FILING DATE: 04-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Noland, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 27866/30822  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 346-5750  
TELEFAX: (312) 984-9740  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 38:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4131 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 148..2910  
PCT-US92-03222-38



Alignment Scores:

Pred. No.:	2.85	Length:	4131
Score:	79.00	Matches:	37
Percent Similarity:	36.00%	Conservative:	17
Best Local Similarity:	24.67%	Mismatches:	50
Query Match:	7.06%	Indels:	46
DB:	5	Gaps:	8

US-10-034-500-2 (1-218) x PCT-US92-03222-38 (1-4131)

```

QY 90 LeuGluIlePheGlyArgLeu-----ThProValIleAspGly 102
    |||||
Db 2133 CTCGAGTAGTGTGTGAGCTCCAGGTTCTTGAGACAGAGTAGACAGAGTGGAGACAGA 2074
    |||||
QY 103 IleIysMeHleGluThrValIleAsnGlnMetProProGlnAlaIleValIleLeu 122
    |||||
Db 2073 AAAGCGTCATCCAGATTGTGTAGGGGATCCG-----GTAGCCCTTCTT 2026
    |||||
QY 123 ThrArgGlyHisAsnAsnLeuGlyMetAsnProValIleGluValIleSerIleTyrGln 142
    |||||
Db 2025 -----CAC-----CATCA 2017
    |||||
QY 143 ThrAsnProAsnValCys-----PheHisValVal 152
    |||||
Db 2016 ACAGAACCGGCGCAGTGTGGGCAATTTGTAGTTATGATGAATTCATGCTCTG 1957
    |||||
QY 153 SerPheAlaAspAspAlaIleGlyIleValIleIleAspGlnIleVal-----AlaLeu 170
    |||||
Db 1956 CAGCATGCTCAGAGATGGCCATGGAAGTGTCTCTGGGCAAGAGCGAGAGTGTATGT 1897
    |||||
QY 171 AsnSerGlySerValLeuValAspGlyLeuGlnLeuGlnAsnProAlaValCysGln 190
    |||||
Db 1896 GAAACTGGCAAGTTGAGATCGATGCGACAGCCACAGGCTGAT-----CCCGCATGAGAG 1843
    |||||
QY 191 GluPhe-----ValAsnSerValPheCysGlnGluGlnIleLeuValThrGln 206
    |||||
Db 1842 AAGTTTGTGTACTCGTCACTCAGAGACCTTCATGTGTGATCATCATCTCATTTGGCAAG 1783
    |||||
QY 207 GluValValIleValArgGlyValAsnPhe 216
    |||||
Db 1782 GTGGTGCATCTACTGCGCTCATCTCTT 1753
    |||||

```

RESULT 29  
US-08-348-891A-1  
; Sequence 1, Application US/08348891A  
; Patent No. 5654136  
; GENERAL INFORMATION:  
; APPLICANT: SASAKI, Keiko  
; APPLICANT: MORI, Takayuki  
; APPLICANT: MAKINO, Satoshi  
; TITLE OF INVENTION: ATTENUATED MEASLES VIRUS VACCINE,  
; TITLE OF INVENTION: CONTAINING SPECIFIC NUCLEOTIDE SEQUENCE AND A METHOD FOR  
; TITLE OF INVENTION: ITS ABSOLUTE IDENTIFICATION  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: YOUNG & THOMPSON  
; STREET: 745 South 23rd Street  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/348,891A  
; FILING DATE: 25-NOV-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/848,400

FILING DATE: 10-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 3-293625  
FILING DATE: 14-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: PATCH, Andrew J.  
REGISTRATION NUMBER: 32,925  
REFERENCE/DOCKET NUMBER: RP-7501  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-521-2297  
TELEFAX: 703-685-0573  
TELEX: 248425 EMBON  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15894 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 108..1682  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1807..3327  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 3438..4442  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 5458..7107  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 7271..9121  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 9234..15782  
US-08-348-891A-1

Alignment Scores:

Pred. No.:	25.5	Length:	15894
Score:	79.00	Matches:	54
Percent Similarity:	34.75%	Conservative:	52
Best Local Similarity:	17.70%	Mismatches:	81
Query Match:	7.06%	Indels:	118
DB:	1	Gaps:	13

US-10-034-500-2 (1-218) x US-08-348-891A-1 (1-15894)

```

QY 7 CysThrIleArgValIleGluSerTyrAsnTyrLeuValAspTyrSerGlySerMetMetMet 26
    |||||
Db 14562 TCCCTTGAAGCCAGGGAGAGCGGCTTGTCTTGGGTGAGGATCGGCTTATGTGATC 14621
    |||||
QY 27 LysHisValAlaValArgGluProIleGluLeuAlaLysGluAlaIleLeuIle 46
    |||||
Db 14622 ACTTAT-----AAGGATTAATAACTA 14645
    |||||
QY 47 AsnAlaAlaMetProIleMetSerTyrGlnIleGlyLeuTyrThr----- 61
    |||||
Db 14646 AAC-----AAGGCTTCTATATAGTGGGGTTCGCCCAATTCTAGATCTGCT 14693
    |||||
QY 62 -----PheAlaProTyrSerValIleIle----- 69
    |||||
Db 14694 CAAAGGAATTAGACACCTTATCCCTCCGAAGTTGGCCTTGTGACACAGATGGAGTA 14753
    |||||
QY 70 -----ProGlnIleSerTyrAsnSerCysVal 78
    |||||
Db 14754 GGTAATATGTCAAAAGTCTTTTAACGGAGGCCGCAAGTACGCTAGCAGTGTGA 14813
    |||||
QY 79 AlaGluCys-----AlaValAsnThrIleLys 87
    |||||
Db 14814 ---GATTGCTCAATTCTATAGTATGAATATCCCTACTTATGTGTGGGTTTATCCAT 14870
    |||||

```

```

Qy 88 SerApleuGlu----- 91
Db 14871 TCAGATATAGAGACCTTCCTAACAAAGATACCTATAGAGAAAGCTAGAGAAATTGGCAGCC 14930
Qy 92 -----IlePheGlyArgLeuThrProVal----- 99
Db 14931 ATCTTATGATGGCTGCTGCTCTGGCAAAATAGATCAATACGTGGATTAAAGCTTATG 14990
Qy 100 -----GlyApeGlyIleLeu-----MetHisGlyThrVal 109
Db 14991 CCTTACAGCGGGGATTGTTGTCAGGATTTATTAAGTATGATGAGGCTTATTTATAGAGAA 15050
Qy 110 IleAenGlnMetProProGln-----AlaIleValIleLeuLeu 122
Db 15051 GTGAACCTTGATACCTCCTAGATACAGCAACTTATCATCTACTGATGATCTTATTTGGTATG 15110
Qy 123 ThrApeGlyHisAenAenLeuGlyMetAenProValGluGluValLeu----- 138
Db 15111 ACAGATCTCAAGCTCAACCGGCTAATGATCTCT---GAAAGATTAAACAGCAGATTAATT 15167
Qy 139 ---SerIleTyGlnThrAenProAenValCySPheHisValIleSerPheHisAenApe 157
Db 15168 GAATCATCTGAGAGACCTTCACCTGACTTATAGCTCACATCTTATCCATTAAGCACTA 15227
Qy 158 AlaGluGlyValAlaIleIleApeGlnIleValAlaLeuAenSer----- 172
Db 15228 AGCTGCATACAGCAATTTGGGAGACGTAAGTTAGTAGAGTATCATCTCTG 15287
Qy 173 -----GlySerValLeuValApeGlyLeuGlnLeuGlnAenPro 186
Db 15288 AAAAACTTACCTATAGACAGCAGGTGCTGATCAATTCGGGTTGGCAGATTAACGAGCT 15347
Qy 187 AlaValCySGlnGluPheValAenSerValPheCySGlnGluIleLeuValThrGlu 206
Db 15348 AAGCTGTCAAAGAAATTATCATCACCATGATGTGCTCCTCAGGCAAGATGATGCTTAAAT 15407
Qy 207 GluValValValLeu 211
Db 15408 TCTATCTCATCTC 15422

RESULT 30
US-08-905-817-1
; Sequence 1, Application US/08905817
; Patent No. 5824777
; GENERAL INFORMATION:
; APPLICANT: SASAKI, Keiko
; APPLICANT: MORI, Takayuki
; APPLICANT: MAKINO, Satoshi
; TITLE OF INVENTION: ATTENUATED MEASLES VIRUS VACCINE.
; TITLE OF INVENTION: CONTAINING SPECIFIC NUCLEOTIDE SEQUENCE AND A METHOD FOR
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: YOUNG & THOMPSON
; STREET: 745 South 23rd Street
; CITY: Arlington
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,817
; FILING DATE: 04-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/348,891
; FILING DATE: 25-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/848,400

```

```

; FILING DATE: 10-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 3-293625
; FILING DATE: 14-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: PATCH, Andrew J.
; REGISTRATION NUMBER: 32,925
; REFERENCE/DOCKET NUMBER: KP-7501A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-521-2297
; TELEFAX: 703-685-0573
; TELEX: 248425 EMBON
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15894 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 108..1682
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1807..3327
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3438..4442
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5458..7107
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 7271..9121
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 9234..15782
; US-08-905-817-1

Alignment Scores:
Pred. No.: 25.5 Length: 15894
Score: 79.00 Matches: 54
Percent Similarity: 34.75% Conservative: 52
Best Local Similarity: 17.70% Mismatches: 81
Query Match: 7.06% Indels: 118
DB: Gaps: 13

US-10-034-500-2 (1-218) x US-08-905-817-1 (1-15894)
Qy 7 CyThrIysArgValGluSerTyAenTyRleuValApeTySerGlySerMetMetMet 26
Db 14562 TGCCTTACGCGAGGAGAGAGCGCTTGTCTTGGGTGAGGATCGGTTCTATGTTATC 14621
Qy 27 LysHisValAlaValArgIleGluLeuAlaIleLeuValIle 46
Db 14622 ACTTAT-----AAGGAGATCTTAAACTA 14645
Qy 47 AsnAlaIleMetProIysMetSerTyGlnGlyLeuTyThr----- 61
Db 14646 AAC-----AGTGCTTCTATATATAGTGGGGTTTCCGCCAATTCTAGATCTGGT 14693
Qy 62 -----PheAlaProTySerValIleIle----- 69
Db 14694 CAAAGGAATTAGACCTTATCCTCCGAACTGGCCTTGTCGAACACAGATGGAGTA 14753
Qy 70 -----ProGlnGlySerTyPAsenSerCyVal 78
Db 14754 GGTATATTTGCAAGTGTCTTTACGCGAGGCCCGAAGTCAGTGGGTAGCAGTGTGA 14813
Qy 79 AlaGluCyS-----AlaValAenThrIleLeu 87
Db 14814 ---GATTGCTTCAATTTCATAGTATGATATATCCCTACCTCTAGTGGGGTTATTCAT 14870

```

```

Qy      88 SerAspLeuGlu----- 91
Db      14871 TCAGATATATAGAGACCTTGCCCTTAACAAGTACTATATAGAGACCTAGAGAAATTGGCAGCC 14930
Qy      92 -----1LePheGlyArgLeuThrProVal----- 99
Db      14931 ATCTTATGATGAGCTCTGCTCTGGGCAAAATAGCATATCTGCTGATTAAGCTTATG 14990
Qy      100 -----GlyAspGlyLeuLys----- MetHisGluThrVal 109
Db      14991 CCTTACAGCGGGGATTTGTGTAGGGATTATTAAGTTATGAGGCTTATTATATAGAGAA 15050
Qy      110 IleAsnGlnMetProGln-----AlaIleValIleLeuLeu 122
Db      15051 GTGAACCTTGATATACCTTATAGATACAGCACTTCAATCTTACTAGATCTTATTGGTTATG 15110
Qy      123 ThrAspGlyHisAsnAsnLeuGlyMetAsnProValGluGluValLys----- 138
Db      15111 ACAGATCTCAAGGCTTAACCGGCTAATGAATCCT---GAAAGATTAAAGCAGACAGATTAAT 15167
Qy      139 ---SerIleTyrGlnThrAsnProAsnValCysPheHisValIleSerPheIleAspAsp 157
Db      15168 GAAATCATCTGTAGAGGACTTACCTGACCTGATATAGTCAATCTTATCACTTAAGCACTA 15227
Qy      158 AlaGluGlyLysAlaIleIleAspGlnIleValAlaLeuAsnSer----- 172
Db      15228 AGCTGATATCAAGCAATTTGTGGAGACGTATTAAGTAGAGGTGATCAATCTTACTCTG 15287
Qy      173 -----GlySerValIleValAspGlyLeuGlnLeuGlnAsnPro 186
Db      15288 AAAAACTTACACCTATAGACGAGGTGATCAATTTGGGGTGGCAATTAAACGACCT 15347
Qy      187 AlaValCysGlnGluPheValAsnSerValPheCysGlnGluGlnIleLeuValThrGlu 206
Db      15348 AAGCTGTGCAGAAATGATTCACATGATGTTCCTCAGGGCAGATGATTCCTTAAT 15407
Qy      207 GluValIleValLeu 211
Db      15408 TCTATCTCATCTCTC 15422

RESULT 31
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.: 3, 17e+05 Length: 4403765
Score: 77.50 Matches: 43
Percent Similarity: 41.04% Conservative: 28
Best Local Similarity: 24.86% Mismatches: 77
Query Match: 6.93% Indels: 25
PB: 4 Gaps: 8

```

```

US-10-034-500-2 (1-218) x US-09-103-840A-2 (1-4403765)
Qy      10 ArgValGluSerTyrAsnTyrLeuValAspTyrSerGlySerMetMetLeuHisVal 29
Db      141768 CGCCAGCAGATACGACGACGTGGTCTGATGTCGCGCGGACGATCGCTCAATG 1417619
Qy      30 AlaValArgGluProLysIleGluLeuAlaLysGluAlaIleLeuLysIleAsnAlaIle 49
Db      1417618 ACCGCGAGGAGAGCT-----TCCGAGGGCCATTCATCAAGTCGCCCACTG 1417571
Qy      50 MetProLysMetSerTyrGlnGlyLeuTyrThrPheAlaProTyrSerValIleIle 69
Db      1417570 ACACCCAGCGCTGCGACCCCGGCTCCATCTACACA---AAGCCGCGCAGCCGACAC 1417514
Qy      70 ProGlnGlySerThrAsnSerCysValAlaGluCys-----Ala 82
Db      1417513 CCGAATCTGGGAATGGGTATACAGCGTGGCCGACATACCAACCTGCTGGCGCATTAACGC 1417454
Qy      83 ValAsnThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGlyAspGly 102
Db      1417453 GTGACCGCCGTTATCACCGAC---GAGGACCAAGCCGACGACCGGATCGCCACCGGC 1417397
Qy      103 IleLysMetHisGluThrValIleAsnGlnMetProGlnAlaIleValIleLeuLeu 122
Db      1417396 ATG-----GCCGCCCTCGACAGATATCTGGCAGCGCGGACGCGGTGATGTC 1417349
Qy      123 ThrAspGlyHisAsnAsnLeuGlyMetAsnProValGluGluValIleSerIleTyrGln 142
Db      1417348 AGCATCAACCGCGATGATCTGGGGCCAGCCCTCGATCAAGAAACCGACAGT-----GCC 1417295
Qy      143 ThrAsnProAsnValCysPheHisValIleSerPheAlaAspAspAlaGluGlyLysAla 162
Db      1417294 GGCAACCGCGGTGTGACCAAGCGGTGTGTGACCGGTGCGATACCGAA----- 1417244
Qy      163 IleIleAspGlnIleValAlaLeuAsn---SerGlySer 174
Db      1417243 -----AACGCGATTTTCATCTCAACGACGACGCGGTAC 1417211

RESULT 32
US-08-470-179-93
; Sequence 93, Application US/08470179
; Patent No. 5645994
; GENERAL INFORMATION:
; APPLICANT: Huang Ph.D, Wei Mun
; TITLE OF INVENTION: Method and Compositions for
; IDENTIFICATION OF SPECIES IN A SAMPLE
; NUMBER OF SEQUENCES: 207
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Trask, Britt and Rossa
; STREET: P. O. Box 2550
; CITY: Salt Lake City
; STATE: Utah
; COUNTRY: USA
; ZIP: 84110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,179
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sweigert Ph.D, Susan E.
; REGISTRATION NUMBER: 36,289
; REFERENCE/DOCKET NUMBER: 2601
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 801-532-1922
; TELEFAX: 801-531-9168
; INFORMATION FOR SEQ ID NO: 93:
; SEQUENCE CHARACTERISTICS:

```

```

1 LENGTH: 423 base pairs
2 TYPE: nucleic acid
3 STRANDEDNESS: double
4 TOPOLOGY: not relevant
5 MOLECULE TYPE: other nucleic acid
6 DESCRIPTION: /desc = "gyra gene segment"
7 HYPOTHETICAL: NO
8 ANTI-SENSE: NO
9 ORIGINAL SOURCE:
10 ORGANISM: Shigella dysenteriae
11 US-08-470-179-93
12
13 Alignment Scores:
14 Pred. No.: 0.133 Length: 423
15 Score: 77.00 Matches: 25
16 Percent Similarity: 40.48% Conservative: 26
17 Best Local Similarity: 19.84% Mismatches: 59
18 Query Match: 6.88% Indels: 16
19 DB: 1 Gaps: 4
20
21 US-10-034-500-2 (1-218) x US-08-470-179-93 (1-423)
22
23 QY TySerValIleIleProGlnGlySerTrpAsnSerCysValAlaGluCysAlaValAsn 84
24   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
25 Db 34 TACGCCATGAACGTACTTAGGCAATGACTGGAACAAAGCCTTATTAATAATCTGCC----- 87
26
27 QY ThrIleIysSerAspLeuGluIlePheGlyArgLeuThrProValIglYAspGlyIleIys 104
28   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
29 Db 88 CGTGCGTGTGTGAC-----GTATCGGTAAATATACATCTCCCATGTGTGACTCG---GCG 138
30
31 QY 105 MethIsgLutThrValIleIleAsnGlnMecProGlnAlaIalValIleLeuLeuThrAsp 124
32   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
33 Db 139 GTTAAATGACAGATCGTCCTGTTGGCGACGACCATTCCTGCTGCATTACATGCTGTAGAC 158
34
35 QY 125 GlyHisAsnAsnLeuGly-----MetAsnPro 133
36   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
37 Db 139 GGTCAAGGTAACCTCGGTTCATGACGCGACGACTCGCGCGCGCAATGCGTTATACGAA 258
38
39 QY 134 ValGluGluValIysSerIleTrpGlnThrAsnProAsnValCysPheHisValIalSer 153
40   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
41 Db 259 ATCCGTCTGGCGAAATATGCCCATGAACCTGATGGCCGATCTCCGAAAAAGACGCTGAT 118
42
43 QY 154 PheAlaAspAspAlaGluGlyIysAlaIleIleAspGlnIleValAlaLeuAsnSerGly 173
44   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
45 Db 319 TTCGTTGATTACTATGACGCGACCGCAAAAAATTCGCCAGCTATGCCAACCAAAATTCCT 378
46
47 QY 174 SerValIleValAspGly 179
48   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
49 Db 379 AACCTGCTGTGAACGCT 396
50
51 RESULT 33
52 US-08-245-511-3
53 Sequence 3, Application US/08245511
54 Patent No. 5928900
55 GENERAL INFORMATION:
56 APPLICANT: Masuure, H Robert
57 APPLICANT: Pearce, Barbara J
58 APPLICANT: Tuomanen, Elaine
59 TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
60 NUMBER OF INVENTION: ACETULAR VACCINES BASED THEREON
61 ADDRESSSEE: 58
62 CORRESPONDENCE ADDRESS:
63 ADDRESSEE: Klauber & Jackson
64 STREET: 411 Hackensack Avenue
65 CITY: Hackensack
66 STATE: New Jersey
67 COUNTRY: USA
68 ZIP: 07601
69 COMPUTER READABLE FORM:
70 MEDIUM TYPE: Floppy disk
71 COMPUTER: IBM PC compatible
72 OPERATING SYSTEM: PC-DOS/MS-DOS
73 SOFTWARE: PatentIn Release #1.0, Version #1.25

```

[illegible]

```

Db 657 CCTGCTTACATGATTAATTAACCTCAGAGAGTCAATCAATGTTGAAGAAAGACAGG 716
Qy 118 aValIleuLeuThrAspGly-----HisAsnLeuGlyMetAsnProValG1 135
Db 717 CTATTAACCTACACAACTGGAGTGTCTACCAATGTAGACCAAGAACTCAAA 776
Qy 135 uGluValIySerIleTyGlnThrAsnProAsnValCysPhe-----His 150
Db 777 ACATCTGGGATATTATTAACAATACAGCAATACCTGCTTCCAGAGATGAATGCA 836
Qy 150 sValIleSerPheAlaAspAspAlaGluIlylAlaIleIleAspGlnIleValAlaLe 170
Db 837 AGTCGCTTACCATGTTGATGTTTCTTAACGTAAGCATTTGCCACGCTAGAGACAG 896
Qy 170 uAsnSerGlySerValIleuValAspGlyLeuGlnLeuLeuGlnAsnProAlaVal 188
Db 897 CCATCAGTCAGTAATGTTTCTTCCCTCGAAT-----AACCAAGCAGTA 939

```

## RESULT 34

```

US-08-600-993A-3
; Sequence 3, Application US/08600993A
; Patent No. 5981229
; GENERAL INFORMATION:
; APPLICANT: Masure, H Robert
; APPLICANT: Pearce, Barbara J
; APPLICANT: Tuomenen, Elaine
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
; TITLE OF INVENTION: ACCELLULAR VACCINES BASED THEREON
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/600,993A
; FILING DATE: 1-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/245,511
; FILING DATE: 18-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/116,541
; FILING DATE: 01-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-069 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521

```

```

; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 960 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; STRAIN: R6

```

```

; IMMEDIATE SOURCE:
; CLONE: SP8U42
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..960
; US-08-600-993A-3

```

```

Alignment Scores:
Pred. No.: 0.503 Length: 960
Score: 77.00 Matches: 48
Percent Similarity: 37.24% Conservative: 41
Best Local Similarity: 20.08% Mismatches: 75
Query Match: 6.88% Indels: 75
DB: Gaps: 9

```

```

US-10-034-500-2 (1-218) x US-08-600-993A-3 (1-960)

```

```

Qy 13 SerTyraenTyfLeuValAspTyf-----SerGlySerMetMetMetLysHis 28
Db 261 AACTTCGACTTCGACCGACACTATTCTCTAAGGCTCAGAGAGCTTGTAAGCATTC 320
Qy 29 ValAlaValArgGlu-ProlySIleGluLeuAlaIleGluAlaIleLeuLysIle----- 46
Db 321 GTTAGACAAAAGCAACCAAGCAAGAAATCTTGACTACTATATTAATTAAGTCTACAT 380
Qy 46 ----- 46
Db 381 GTCTAATGGAACTATGGAATGCAGACAGCTCAAAACTATGTTGTAAGACCTCAA 440
Qy 47 -AsnAlaIleMetProLysMetSerTyfGlnGlyIleu-----TyfThrPh 62
Db 441 TATATTGAATTATTAACCTCAAGTTAGCTTGGCTGGAATGCTTCAGCAACCAACCAATA 500
Qy 62 eAlaProTyfSerValIleIleProGlnGlySerTyfAsnSerCysValAlaGluCysAl 82
Db 501 TGACCCCTATTGA-----CATCCAGAGAGAGCCCAAGACCG 536
Qy 82 aValAsnThrIleLysSerAspLeuGlnIlePheGlyArgLeu----- 96
Db 537 CGGAACCTTGCTTATCTGAATGAAATCAAGCTACATCTCTGCTGAACAGTATGA 596
Qy 97 -----ThrProValGlyAspGlyIlyIleLysMetHisGluThr----- 108
Db 597 GAAGAGCTCAATACACCAATTACTGATGGCTCAAAAGTCTCAATCAAGCAAGTAATTA 656
Qy 109 -----ValIleAsnGlnMetProProGlnAlaAl 118
Db 657 CCTGCTTACATGATTAATTAACCTCAGAGAGTCAATCAATGTTGAAGAAAGACAGG 716
Qy 118 aValIleuLeuThrAspGly-----HisAsnLeuGlyMetAsnProValG1 135
Db 717 CTATTAACCTACACAACTGGAGTGTCTACCAATGTAGACCAAGAACTCAAA 776
Qy 135 uGluValIySerIleTyGlnThrAsnProAsnValCysPhe-----His 150
Db 777 ACATCTGGGATATTATTAACAATACAGCAATACCTGCTTCCAGAGATGAATGCA 836
Qy 150 sValIleSerPheAlaAspAspAlaGluIlylAlaIleIleAspGlnIleValAlaLe 170
Db 837 AGTCGCTTACCATGTTGATGTTTCTTAACGTAAGCATTTGCCACGCTAGAGACAG 896
Qy 170 uAsnSerGlySerValIleuValAspGlyLeuGlnLeuLeuGlnAsnProAlaVal 188
Db 897 CCATCAGTCAGTAATGTTTCTTCCCTCGAAT-----AACCAAGCAGTA 939

```

## RESULT 35

```

US-08-961-083-1
; Sequence 1, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452

```



```

FEATURE:
NAME/KEY: mat peptide
LOCATION: 1..2046
US-08-481-435-5

Alignment Scores:
Pred. No.: 1.73 Length: 2049
Score: 77.00 Matches: 48
Percent Similarity: 37.24% Conservative: 41
Best Local Similarity: 20.08% Mismatches: 75
Query Match: 6.88% Indels: 75
DB: Gaps: 9

US-10-034-500-2 (1-218) x US-08-481-435-5 (1-2049)

QY 13 SerTyraAsnTyrlEuaValAspTYr-----SerglySerMetMetWellyshis 28
DB 297 AACTTCGACTCGACCAACCACTATTTCTCGTAAGGCTCGAAGCTTGTTAGCGATTCA 356
QY 29 ValAlaValArgGlu-ProlysiLeGluEuaLysGluAlaLeuLysile-----46
DB 357 GTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAAATAGGTCTACAT 416
QY 46 -----46
DB 417 GTCTAATGGGAAGTATGATGACAGACAGAGCTCAAACTACTATGTAAAGACTCAA 476
QY 47 -AsnAlaAlaMetProlyseMetSerTyrlnglyleu-----TyrThrph 62
DB 477 TAATTTAAGTTAAGCTTCACTGACCTTGCTGCTGGAATGACCTCAGCAACCAACATA 536
QY 62 eAlaProTySerVallelleProGlnGlySerTyraSerCyValAlaGluCysAl 82
DB 537 TGACCCCTTATCA-----CATCCAGAGCAGACCCCAACCG 572
QY 82 aValaenThrilleYserAspLeuGluillePheGlyArgleu-----96
DB 573 CGGAACCTTGCTTATCTGGAATGAAATCAAGCTCATCTCTGTGAACAGTATGA 632
QY 97 -----ThrProValGlyAspGlylleYsMetHieGluThr-----108
DB 633 GAAAGCAATCAATACCAATTAAGTGGCTCAAAAGTCTCAACATCGCAAGATATTA 692
QY 109 -----ValleAsnGlnMetProProGlnAlaAl 118
DB 693 CCCTGCTTACATGATTAATTAAGTCAAGAGATCATCATCAAGTTGAAGAAACAG 752
QY 118 aValleleuLeuThrAspGly-----HisAsnAsnLeuGlyMetAsnProValG 135
DB 753 CTATAACCTACTCACAACGAGATGATGCTACACAAATGTAGACCAAGAACTCAAA 812
QY 135 uGluVallylSerilleTyrlGlnThrAsnProAsnValCysPhe-----Hi 150
DB 813 ACATCTGTGGATATTATTAATACAGACGAATACGTTGCCCTACAGACGATGAATGA 872
QY 150 eValValSerPheAlaAspAspAlaGluGlyLysAlaIlelleAspGlnilleValAla 170
DB 873 AGTCGCTTACCATGTTGATGTTCTTAACGTAAGATATTGCCAGCTAGAGACAG 932
QY 170 uAsnSerGlySerVallelleProGlnGlySerTyraSerCyValAlaGluCysAl 188
DB 933 CCATCACTCAAGTAAATGTTCTTCCGAAT-----AACCAAGCACTA 975

RESULT 37
US-08-961-527-145/c
Sequence 145, Application US/08961527
GENERAL INFORMATION:
APPLICANT: Charles Kunesch
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESS:
ADDRESSER: Human Genome Sciences, Inc.

```

```

STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 145:
SEQUENCE CHARACTERISTICS:
LENGTH: 10711 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

US-08-961-527-145

Alignment Scores:
Pred. No.: 25.6 Length: 10711
Score: 77.00 Matches: 48
Percent Similarity: 37.24% Conservative: 41
Best Local Similarity: 20.08% Mismatches: 75
Query Match: 6.88% Indels: 75
DB: Gaps: 9

US-10-034-500-2 (1-218) x US-08-961-527-145 (1-10711)

QY 13 SerTyraAsnTyrlEuaValAspTYr-----SerglySerMetMetWellyshis 28
DB 9518 AACTTCGACTCGACCAACCACTATTTCTCGTAAGGCTCGAAGCTTGTTAGCGATTCA 9459
QY 29 ValAlaValArgGlu-ProlysiLeGluEuaLysGluAlaLeuLysile-----46
DB 9458 GTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAAATAGGTCTACAT 9399
QY 46 -----46
DB 9398 GTCTAATGGGAAGTATGATGACAGACAGCTCAAACTACTATGTGAAGACTCAA 9339
QY 47 -AsnAlaAlaMetProlyseMetSerTyrlnglyleu-----TyrThrph 62
DB 9338 TAATTTAAGTTAAGCTTCACTGACCTTGCTGCTGGAATGACCTCAGCAACCAACATA 9279
QY 62 eAlaProTySerVallelleProGlnGlySerTyraSerCyValAlaGluCysAl 82
DB 9278 TGACCCCTTATCA-----CATCCAGAGCAGACCCCAACCG 9243
QY 82 aValaenThrilleYserAspLeuGluillePheGlyArgleu-----96
DB 9242 CGGAACCTTGCTTATCTGGAATGAAATCAAGCTCATCTCTGCTGAACAGTATGA 9183
QY 97 -----ThrProValGlyAspGlylleYsMetHieGluThr-----108
DB 9182 GAAAGCAATCAATACCAATTAAGTCAAGATCAATCAATCAAGTGAAGAAACAG 9123
QY 109 -----ValleAsnGlnMetProProGlnAlaAl 118
DB 9122 CCCTGCTTACATGATTAATTAAGTCAAGAGATCATCATCAAGTTGAAGAAACAG 9063

```

```

Qy 118 aValIleuLeuThrAspGly-----HisAsnAsnLeuGlyMetAsnProValGI 135
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9062 CTAACTACTACTCACTGCGATGATGCTACACAAATGAGCAGCAAGACTCAAAA 9003
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 135 uGluValIleSerIleTyGlnThrAsnProAsnValCysphe-----HI 150
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9002 ACACCTGTGGGATATTTCACATACACAGCAATACCTGCTTACAGACGATGAATTGCA 8943
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 150 aValIleSerPheIleAspAspAlaGluGlyLysAlaIleIleAspGlnIleValAlaIle 170
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8942 AGTCGCTTCAACCATGTTGATGTTTCTTACACGGTAAGTCATGCTCCAGCTAGAGCAGC 8883
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 170 uAsnSerGlySerValIleuValAspGlyLeuGlnLeuLeuGlnAsnProAlaVal 188
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8882 CCATCAGTCAGATGATGTTTCCTTCGGAAT-----AACCAAGACAGTA 8840
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 38
US-08-936-165A-199
; Sequence 199, Application US/08936165A
; Patent No. 6348582
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Burnham, Martin
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Lonetto, Michael
; APPLICANT: Nicholas, Richard
; APPLICANT: Pratt, Julie
; APPLICANT: Reichard, Richard
; APPLICANT: Rosenberg, Martin
; APPLICANT: Ward, Judith
; TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,
; TITLE OF INVENTION: Polypeptides and Their Uses
; NUMBER OF SEQUENCES: 534
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,165A
; FILING DATE: 24-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/027,032
; FILING DATE: 24-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimm, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50549
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 199:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2178 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; US-08-936-165A-199

Alignment Scores: 2.25 Length: 2178
Pred. No.:

```

```

Score: 76.50 Matches: 47
Percent Similarity: 37.38%
Best Local Similarity: 22.82%
Query Match: 6.84%
DB: 4 Gaps: 9

US-10-034-500-2 (1-218) x US-08-936-165A-199 (1-2178)
Qy 17 LeuValAspTySerGlySerMetMetLysHisValAlaValArgGluPro----- 34
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1307 CTAAATGAAAGAGCATTCAATCATGTGGCATATTTAAATATTCATACGGCTTATGAT 1366
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 35 -----LysIleGluLeuAlaLysGluAlaIleLeuLysIleAsnAla 48
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1367 TTGTTAAATTCAGCTTAAATAATGAAAGATCCGTAAGACTTGCTGTCTGAAAATGTT 1426
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 49 AlameTProLysMetSerTyGlnGlyLeuTyThrPheAlaProTySerValIle 68
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1427 GATGCACCTTGCCATACACAGACCAATGATGTATGTGTTCTTAAATTT----- 1477
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 69 IleProGlnGlySerTyPheAsnSerCysValAlaGluCysAlaValAsnThrIleLysSer 88
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1478 -----TATGATCATGTATACCA-----AATTAATTTAA----- 1507
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 89 AspLeuGlnIlePheGlyArgLeuThrProValGlyAspGlyIleLysMetHisGluThr 108
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1508 -----CCGATTTTGGTATACAAATATATGTGACAAATGATTAATAATACAGTGAAACA 1561
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 109 ValIleAsnGlnMetProProGlnAlaAlaValIleLeuLeuThrAspGlyHisAsnAsn 128
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1562 GTTGTTCTA-----GCTTAAATAATAT 1582
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 129 LeuGlyMetAsnProValGluGluValLysSerIleTyGln----- 142
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1583 GATGATTA-----AAAGATTGTATCACTATCATCGCAATAATAAA 1624
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 143 -----ThrAsnProAsnValCysPheHis-----Val 151
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1625 ATGAACAATGAAAGAAATGTTCACTTTGCACTTTCACAAATTTTCGAAATTTGATT 1684
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 152 ValSerPheAlaAspAspAlaGluGlyLysAlaIleIleAspGlnIle----- 167
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1685 ATCATTTTAAATAATGTTGCAGACGACATCGTACATTTGTTCAAGTTGTGATTGCGAT 1744
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 168 -----ValAlaLeuAsnSerGlySerValIleuValAspGlyLeuGlnLeuGlnAsn 185
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1745 GAAGATACGATTTTGAATCATCAAAAGTTTATGTTTACGGGTATTAAGCACGATATGATT 1804
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 186 ProAlaValCysGlnGlu 191
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1805 CAAATGTTTCTTACCAA 1822
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 39
US-09-193-562D-27
; Sequence 27, Application US/09193562D
; Patent No. 6309857
; GENERAL INFORMATION:
; APPLICANT: Pauli, Benedicht U.
; TITLE OF INVENTION: Nucleotide Sequences Encoding Mammalian Calcium
; TITLE OF INVENTION: Activated Chloride Channel-Adhesion Molecules
; FILE REFERENCE: 18617.0052
; CURRENT APPLICATION NUMBER: US/09/193,562D
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: US/60/065,922
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 27
; LENGTH: 3007
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-193-562D-27

Alignment Scores:

```



Pred. No.:	3.8	length:	3007
Score:	76.50	Matches:	42
Percent Similarity:	37.26%	Conservative:	37
Best Local Similarity:	19.81%	Mismatches:	82
Query Match:	6.84%	Indels:	51
DB:	4	Gaps:	8

Oy	38	LeuAlaIysGluValaIleLeuValIleAsnAlaIleAspProIysMetSerYrGlnGly	57
Db	977	CTTGACAAATCTGGAGACATGGACACTGTGTAACCCGCTCAATGACTGAATCAAGACGC	1036
Oy	58	GlyLeuYrThrRhpAlaProYrSerValIleIleProGlnIYserTrpAsnSerCys	77
Db	1037	CAGCTTTTCTGTGTGCAG-----ACAGTTGAGCTGGGGTCTGGGTGGGATG	1084
Oy	78	ValAlaGluCysAlaValaAsnThrIleIleYserTrpAspLeu	90
Db	1085	GTGACATTGACAGCTGTGCTGCCATGTGCAAAAGTAACCTCATACAGTAAACAGTGGCAGT	1144
Oy	91	-----GluIlePheGlyArgLeu-----ThrProValGly	100
Db	1145	GACAGGGACACACATCTGCCAAAGTTTACTGTGCAGACAGCTTCAGAGAGGACGTCATCTGC	1204
Oy	101	AspGlyIleYrMetHisGluThrValIleAsnGlnMetProGln-----AlaAla	118
Db	1205	AGCGGGCTTTCATCGCGCATTTTACTGTATTAGGAGAAATATCAACATGATGATCTGTA	1264
Oy	119	ValIleLeuLeuThrAspGlyHisAsnAsp-----	128
Db	1285	ATTGTGCTGTGACGGATGGGAAAGACACACTTTAAGTGGGTGCTTTAACAGAGCTCAA	1324
Oy	129	-----LeuGlnMetAsnProValGluGluValIys	138
Db	1325	CAAAAGTGTGCATCATCCACACAGTGGCTTTGGGGCCCTCTGCAGCTCAAGAACTAAG	1384
Oy	139	SerIleYrGlnThrAsnProAsnValCysPheHisValValSerPheAlaAspAla	158
Db	1385	GAGCTGTCCAAATGACAGAGGTTTACAGACATATCT-----TCAGATCAAGTT	1435
Oy	159	GluGlyIysAlaIleIleAspGlnIleValAlaLeuAsnSerGlySerValLeuValaAsp	178
Db	1436	CAGAACATGGCCCTCATTTGATGCTTTGGGGCCCTTTCATCAGAGAAATGAGCTGTCT	1495
Oy	179	-----GlyLeuGlnIleLeuGlnAsnProAlaValCysGlnGlnPheValaAsnSerVal	196
Db	1486	CAGCCCTCATCTCAGCTTGAGATGAGGATTAACCTCCAG-----AACAGCCAG	1546
Oy	197	PheCysGlnGluGlnIleLeuValThrGluGluVal	208
Db	1547	TGGATGAATGGCACAGTATGTGTGACACAGCCGTG	1582

RESULT 40  
US-09-245-2488-54  
Sequence 54, Application US/09245248B  
Patent No. 6395472  
GENERAL INFORMATION:  
APPLICANT: Abbott Laboratories  
APPLICANT: Leary, Thomas  
APPLICANT: Eker, James  
APPLICANT: Chalmers, Michelle  
APPLICANT: Simons, John  
APPLICANT: Birkemeyer, Larry  
APPLICANT: Muerhoff, Scott  
APPLICANT: Pilot-Matias, Tami  
APPLICANT: Desai, Suresh  
APPLICANT: Mushahwar, Isa  
TITLE OF INVENTION: METHODS OF UTILIZING THE TT VIRUS  
FILE REFERENCE: 6461, US, 01  
CURRENT APPLICATION NUMBER: US/09/245,248B  
CURRENT FILING DATE: 1999-02-05  
NUMBER OF SEQ ID NOS: 71

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 54
;
; LENGTH: 3839
;
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-245-248B-54

```

Alignment Scores:	
Pred. No.:	5.65
Score:	76.50
Percent Similarity:	38.40%
Best Local Similarity:	24.00%
Query Match:	6.84%
DB:	4
	Gaps:
US-10-034-500-2 (1-218) × US-09-245-246B-54 (1-3839)	4

QY	34	ProlyslIleGluMetAlaLysGluValIleIleuLysIleAsnAlaIleMetProLysMet	53
Db	1269	CCCCAGACAGACCTCTGTGTACATGACACATCTCTTACGATCTTTGGCTCTGGCGGGCATATG	1328
QY	54	serTyrgInglyIgy-----LeuTyrrhrpheAlaProTySerValIleIleProGln	71
Db	1329	CAATTCGCGTCCGGCTCCACCATTAACGTACACTATGATGGTGTGTCAATCCCAAGATCTCGAA	1388
QY	72	GlySerTrpAsnSerCysValAlaIleGluCysAlaValAsnThrIleLysSerAspLeuGlu	91
Db	1389	TCCATGTGTCACACGCTCCGAGGTGTACTTCTACTAACTTTACAGAAACATCA-----	1442
QY	92	IlephedGlyArgLeuThrProValIleAspGlyIleLeuMetHisGluThrValIleAsn	111
Db	1443	-----GGCAAGGCGCACATTTACATGATATAATTATTAACT	1478
QY	112	GlnMetPro-----ProGlnAlaIleValIleLeuLeuThrAspGly	125
Db	1479	CATTAACTCTACTACACACACACACAAACCAAGACCAATTTAAAGATTTTACAGAAAAAC	1538
QY	126	His-----AsnAsnLeuGlyMetAsnProValGluGluValIlysserIle	140
Db	1539	CAAGAGCAACAAATGGAAACAAATGTATGGCGAAACCTAAGTAAACAGCTGTAAATTTAAAC	1598
QY	141	TyrGlnThrAsnPro	145
Db	1599	AAACAGATACCTT	1613

Search completed: April 8, 2003, 06:19:25  
Job time : 910 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 8, 2003, 05:47:04 ; Search time 97 Seconds  
(without alignments)  
1971.367 Million cell updates/sec

Title: US-10-034-500-2

Perfect score: 1119  
Sequence: 1 AEVTASCTKVESYNYLVY.....QEQILVTEBVVLGVNFAF 218

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Delop 6.0 , Delext 7.0

Searched: 593429 seqs, 438583890 residues

Total number of hits satisfying chosen parameters: 1186858

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q/cgnt2\_1/USPRO.apool/US10034500/runat 02042003 092747 19820/app\_query.fasta.1.391  
-DB=Published Applications NA -QFWT=fastlap -SUFFIX=mpb -MINMATCH=0.1  
-LOOPC=0 -LOOEXT=0 -UNITS=dices -START=1 -END=1 -MATRIX=blomsum62  
-TRANS=human40.cdi -LIST=45 -DOCCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=40 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAILEN=200000000 -USER=US10034500 @CGN 1.1.69 @runat 02042003 092747 19820  
-NCPU=6 -ICPU=3 -NO KLPXY -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications NA:

- 1: /cgnt2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgnt2\_6/ptodata/1/pubpna/PCR\_NEW\_PUB.seq:\*
- 3: /cgnt2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgnt2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgnt2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgnt2\_6/ptodata/1/pubpna/PTUTS\_PUBCOMB.seq:\*
- 7: /cgnt2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgnt2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgnt2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgnt2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgnt2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgnt2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgnt2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgnt2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	93.5	8.4	687	10 US-09-974-300-2433	Sequence 2433, Ap
2	91.5	8.2	3043	12 US-10-025-167-16	Sequence 16, Appl
3	91.5	8.2	3169	9 US-09-981-353-53	Sequence 53, Appl
4	91.5	8.2	3181	12 US-10-025-167-18	Sequence 18, Appl

5	91.5	8.2	3265	9 US-09-992-598-378	Sequence 378, App
6	91.5	8.2	3265	9 US-09-989-2934-378	Sequence 378, App
7	91.5	8.2	3265	9 US-10-063-547-69	Sequence 69, Appl
8	91.5	8.2	3265	9 US-09-989-735-378	Sequence 378, App
9	91.5	8.2	3265	9 US-09-990-444-378	Sequence 378, App
10	91.5	8.2	3265	9 US-09-989-730-378	Sequence 378, App
11	91.5	8.2	3265	9 US-09-990-436-378	Sequence 378, App
12	91.5	8.2	3265	9 US-09-991-181-378	Sequence 378, App
13	91.5	8.2	3265	9 US-09-993-687-378	Sequence 378, App
14	91.5	8.2	3265	9 US-09-989-734-378	Sequence 378, App
15	91.5	8.2	3265	9 US-09-997-653-378	Sequence 378, App
16	91.5	8.2	3265	9 US-10-174-590-257	Sequence 257, App
17	91.5	8.2	3265	9 US-10-176-758-257	Sequence 257, App
18	91.5	8.2	3265	9 US-10-063-616-69	Sequence 69, Appl
19	91.5	8.2	3265	9 US-10-175-737-257	Sequence 257, App
20	91.5	8.2	3265	9 US-09-993-667-378	Sequence 378, App
21	91.5	8.2	3265	9 US-10-063-502-69	Sequence 69, Appl
22	91.5	8.2	3265	9 US-10-173-706-257	Sequence 257, App
23	91.5	8.2	3265	9 US-10-175-738-257	Sequence 257, App
24	91.5	8.2	3265	9 US-10-175-752-257	Sequence 257, App
25	91.5	8.2	3265	9 US-10-176-482-257	Sequence 257, App
26	91.5	8.2	3265	9 US-10-176-757-257	Sequence 257, App
27	91.5	8.2	3265	9 US-10-176-913-257	Sequence 257, App
28	91.5	8.2	3265	9 US-10-180-552-257	Sequence 257, App
29	91.5	8.2	3265	9 US-10-180-557-257	Sequence 257, App
30	91.5	8.2	3265	9 US-09-990-438-378	Sequence 378, App
31	91.5	8.2	3265	9 US-09-990-562-378	Sequence 378, App
32	91.5	8.2	3265	9 US-09-997-428-378	Sequence 378, App
33	91.5	8.2	3265	9 US-09-997-666-378	Sequence 378, App
34	91.5	8.2	3265	9 US-10-173-700-257	Sequence 257, App
35	91.5	8.2	3265	9 US-10-174-572-257	Sequence 257, App
36	91.5	8.2	3265	9 US-10-174-579-257	Sequence 257, App
37	91.5	8.2	3265	9 US-10-174-588-257	Sequence 257, App
38	91.5	8.2	3265	9 US-10-175-739-257	Sequence 257, App
39	91.5	8.2	3265	9 US-10-175-740-257	Sequence 257, App
40	91.5	8.2	3265	9 US-10-175-743-257	Sequence 257, App
41	91.5	8.2	3265	9 US-10-176-488-257	Sequence 257, App
42	91.5	8.2	3265	9 US-10-176-492-257	Sequence 257, App
43	91.5	8.2	3265	9 US-10-176-747-257	Sequence 257, App
44	91.5	8.2	3265	9 US-10-176-750-257	Sequence 257, App
45	91.5	8.2	3265	9 US-10-176-750-257	Sequence 257, App

ALIGNMENTS

RESULT 1  
US-09-974-300-2433  
; Sequence 2433, Application US/09974300  
; Patent No. US20020146721A1  
; GENERAL INFORMATION:  
; APPLICANT: Berka, Randy M.  
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene  
; FILE REFERENCE: 10085,500-US  
; CURRENT APPLICATION NUMBER: US/09/974,300  
; PRIOR APPLICATION NUMBER: 2001-10-05  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 09/680,598  
; PRIOR FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 8481  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2433  
; LENGTH: 687  
; TYPE: DNA  
; ORGANISM: Bacillus licheniformis  
US-09-974-300-2433

Alignment Scores:  
Pred. No.: 0.00384  
Score: 93.50  
Percent Similarity: 34.43%

Length: 687  
Matches: 46  
Conservative: 27



Db 1277 TGTATTGATGAGTGA---CAAGTGGGCCATTGTTTATTGCTTTG 1327  
Qy 153 Serphea1-----155  
Db 1328 GGAAGAGTGTGCTGATGAGCACTAATAGATGACGAATACAGGAGGAGTCAATTT 1387  
Qy 156 -----Aapapalagluclylysalalalleleapgnillevalalaleuanser 172  
Db 1388 TATGTTTCAGATGAGTCAAGACATGCGCTCATTTGTTGGGCTCTTACATCA 1447  
Qy 173 GlySer 174  
Db 1448 GGAAT 1453

RESULT 3  
US-09-981-353-53  
Sequence 53. Application US/09981353  
Patent No. US20020160382A1  
GENERAL INFORMATION:  
APPLICANT: Laeek, Amy W.  
APPLICANT: Jones, David A.  
TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER  
FILE REFERENCE: PA-0038 US  
CURRENT FILING DATE: 2001-10-11  
CURRENT APPLICATION NUMBER: US/09/981,353  
NUMBER OF SEQ ID NOS: 194  
SOFTWARE: PERL Program  
SEQ ID NO 53  
LENGTH: 3169  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. US20020160382A1 2771481GB1  
US-09-981-353-53

Alignment Scores:  
Pred. No.: 0.077 Length: 3169  
Score: 91.50 Matches: 46  
Percent Similarity: 35.64% Conservative: 26  
Best Local Similarity: 22.77% Mismatches: 55  
Query Match: 8.18% Indels: 75  
DB: 9 Gaps: 8  
US-10-034-500-2 (1-218) x US-09-981-353-53 (1-3169)  
Qy 17 LeuValaApTyTSeRgLySeMeTMeTcLyShIsvAlaValaArgLupProLyStle 36  
Db 952 GTTCTGTATAGTCTGGAAGCATGGG-----978  
Qy 37 GluLeuAlaYsGluAlaIleLeuValleleuAlaIleMetProLyMeSeSeTyGln 56  
Db 979 -----GGTAAGACCGCTTAATGATGATCAAGCAAGCAAAATTTCTGCTGAG 1032  
Qy 57 GlyGlyLeuTyTThPheAlaProTySeRValleleleProGlnGlySeTyPAsnSer 76  
Db 1033 -----ACTGTGAATAATGATCTCGGTGGG 1059  
Qy 77 CysVal-----AlaGlu 80  
Db 1060 ATGTTGACTTGTGATGACTGCCACTATTGTAATAAGCTAATCAATAAAACAGT 1119  
Qy 81 CysAlaValaAntThIleYsSerApleuGluIlePheGlyArgLeuThProValGly 100  
Db 1120 GATGAAAGAACACACTATGCGAGATTACCTCATAT-----CCTGTGGGA 1167  
Qy 101 -----AapGlyIleYsMeThIsGluThValleleuGlnMeSeProPro 115  
Db 1168 GGAATTCATCTGCTGCAATTAAATGATTCAGGGATTGAGAGCTACATTC 1227  
Qy 116 Gln-----AlaAlaValaIleleuLeuThAapGlyHIsaenAsnLeuGlyMeTAsn 132  
Db 1228 CAATCGATGATCGAAGTACTGCTGCTGATGATGGAGGATTAACATGCAAGTTCT 1287

Qy 133 ProValGluGluVallySeSerIleTyGlnThAsnProAsnValCysPheHIsvala 152  
Db 1288 TGTATTGATGAGTGA---CAAGTGGGCCATTGTTTATTGCTTTG 1338  
Qy 153 Serphea1-----155  
Db 1339 GGAAGAGTGTGCTGATGAGCACTAATAGATGACGAATACAGGAGGAGTCAATTT 1398  
Qy 156 -----Aapapalagluclylysalalalleleapgnillevalalaleuanser 172  
Db 1399 TATGTTTCAGATGAGTCAAGACATGCGCTCATTTGTTGGGCTCTTACATCA 1458  
Qy 173 GlySer 174  
Db 1459 GGAAT 1464

RESULT 4  
US-10-025-167-18  
Sequence 18. Application US/10025167  
Patent No. US20020127693A1  
GENERAL INFORMATION:  
APPLICANT: BILLING-MEDEL, PATRICIA A.  
COHEN, MAURICE  
COLPITTS, TRACEY L.  
FRIEDMAN, PAULA N.  
HAYDEN, MARK  
KLASS, MICHAEL R.  
ROBERTS-RAPP, LISA  
RUSSELL, JOHN C.  
STROUPE, STEPHEN D.  
TITLE OF INVENTION: REAGENTS AND METHODS FOR THE  
USEFUL FOR DETECTING DISEASES OF THE GASTROINTESTINAL  
TRACT  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Abbott Laboratories  
STREET: 100 Abbott Park Road  
CITY: Abbott Park  
STATE: IL  
COUNTRY: USA  
ZIP: 60064-3500  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: PASCSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/025,167  
FILING DATE: 19-Dec-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/049,698  
FILING DATE: <Unknown>  
APPLICATION NUMBER: 08/828,856  
FILING DATE: 31-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Becker, Cheryl L.  
REGISTRATION NUMBER: 35,441  
REFERENCE/DOCKET NUMBER: 6068, US, P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 847/935-1729  
TELEFAX: 847/938-2623  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3181 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 18:  
US-10-025-167-18

**Alignment Scores:**

Pred. No.:	0.0774	Length:	3181
Score:	91.50	Matches:	46
Percent Similarity:	35.64%	Conservative:	26
Best Local Similarity:	22.77%	Mismatches:	55
Query Match:	8.18%	Indels:	75
DB:	12	Gaps:	8

US-10-034-500-2 (1-218) X US-10-025-167-18 (1-3181)

```

Oy 17 LeuValAspTrpSerGlySerMetMetMetLeuHisIleValAlaValAlaGluProValIle 36
Db 952 GTTCTGCTGAAGTCTGGAAAGCATGGG----- 978
Oy 37 GluLeuAlaLysGluValAlaIleLeuLysIleAsnAlaAlaMetProLysMetSerTyGln 56
Db 979 -----GGTAAGGACCCCGCTAAATCGAATGAATCAAGACGACAAACATTCTCGTGGAC 1032
Oy 57 GlyGlyLeuTyThrPheAlaProTySerValIleIleProGlnGlySerTrpAsnSer 76
Db 1033 -----ACGTGTGAATAATGAGATCGGGGTGGG 1059
Oy 77 CysVal-----AlaGlu 80
Db 1060 ATGGTTCACTTGTGATGACTGCCACTATTGTAAATGAAGCTAAATCAATAAAGACAT 1119
Oy 81 CysAlaValAsnThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGly 100
Db 1120 GATGAAGAAGAACACACATCGACGAGATTTCACACATAT-----CTCTGGGA 1167
Oy 101 -----AspGlyIleLysMetHisGluThrValIleAsnGlnMetProPro 115
Db 1168 GAAACTTCACATCGCTCGCGAATTAAATATGCAATTCAGGTGATGTGAAGACTACATTC 1227
Oy 116 Gln-----AlaAlaValAlaIleLeuLeuThrAspGluHisAsnAsnLeuLysMetAsn 132
Db 1228 CAATCCGATGGATCGAAGATGCTGCTCTGACTGATGGAGGAGATTAACCTGCAAGTTCT 1287
Oy 133 ProValGluGluValLysSerIleTyGlnThrAsnProAsnValCysPheHisValVal 152
Db 1288 TGTATTATATGAAGTGAAA-----CAAGGGGGCCATTGTTTATTATTCGCTTTG 1338
Oy 153 SerPheAla----- 155
Db 1339 GAAAGAGCTGCTGATGGAAGCATTAATAGATGAGACAGATTAACAGAGGAATCATTTT 1398
Oy 156 -----AspAspAlaGluGlyLysAlaIleIleAspGlnIleValAlaIleAsnSer 172
Db 1339 TATGTTTCAGATGAAGCTCAGAACATAGCGCTCATGTGATGCTTTTGGGGCTCTTACATCA 1458
Oy 173 GlySer 174
Db 1459 GGAAT 1464

RESULT 5
US-09-992-598-378
; Sequence 378, Application US/0992598
; Patent No. US20020160384A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Deenoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Garritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.

```

APPLICANT: Napier, Mary A.  
 APPLICANT: Pan, James  
 APPLICANT: Paoni, Nicholas F.  
 APPLICANT: Roy, Margaret Ann  
 APPLICANT: Stewart, Timothy A.  
 APPLICANT: Tumas, Daniel  
 APPLICANT: Watanabe, Colin K.  
 APPLICANT: Williams, P. Mickey  
 APPLICANT: Wood, William I.  
 APPLICANT: Zhang, Zemin  
 TITLE OF INVENTION: Acids and Transmembrane Polypeptides and Nucleic  
 TITLE OF INVENTION: Acids Encoding the Same  
 FILE REFERENCE: P2730PIC20  
 CURRENT APPLICATION NUMBER: US/09/992,598  
 CURRENT FILING DATE: 2001-11-14  
 PRIOR APPLICATION NUMBER: 60/049787  
 PRIOR FILING DATE: 1997-06-16  
 PRIOR APPLICATION NUMBER: 60/062250  
 PRIOR FILING DATE: 1997-10-17  
 PRIOR APPLICATION NUMBER: 60/065186  
 PRIOR FILING DATE: 1997-11-12  
 PRIOR APPLICATION NUMBER: 60/065311  
 PRIOR FILING DATE: 1997-11-13  
 PRIOR APPLICATION NUMBER: 60/066770  
 PRIOR FILING DATE: 1997-11-24  
 PRIOR APPLICATION NUMBER: 60/075945  
 PRIOR FILING DATE: 1998-02-25  
 PRIOR APPLICATION NUMBER: 60/078910  
 PRIOR FILING DATE: 1998-03-20  
 PRIOR APPLICATION NUMBER: 60/083322  
 PRIOR FILING DATE: 1998-04-28  
 PRIOR APPLICATION NUMBER: 60/084600  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/087106  
 PRIOR FILING DATE: 1998-05-28  
 PRIOR APPLICATION NUMBER: 60/087607  
 PRIOR FILING DATE: 1998-06-02  
 PRIOR APPLICATION NUMBER: 60/087609  
 PRIOR FILING DATE: 1998-06-02  
 PRIOR APPLICATION NUMBER: 60/087759  
 PRIOR FILING DATE: 1998-06-02  
 PRIOR APPLICATION NUMBER: 60/087827  
 PRIOR FILING DATE: 1998-06-03  
 PRIOR APPLICATION NUMBER: 60/088021  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088025  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088026  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088028  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088029  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088030  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088033  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088326  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088167  
 PRIOR FILING DATE: 1998-06-05  
 PRIOR APPLICATION NUMBER: 60/088202  
 PRIOR FILING DATE: 1998-06-05  
 PRIOR APPLICATION NUMBER: 60/088212  
 PRIOR FILING DATE: 1998-06-05  
 PRIOR APPLICATION NUMBER: 60/088217  
 PRIOR FILING DATE: 1998-06-05  
 PRIOR APPLICATION NUMBER: 60/088655  
 PRIOR FILING DATE: 1998-06-09  
 PRIOR APPLICATION NUMBER: 60/088734  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088738  
 PRIOR FILING DATE: 1998-06-10

PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090431  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090435  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090444  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090472  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090535  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090540  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090542

PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090557  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090676  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090678  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090690  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090694  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090695  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090696  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090862  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/090863  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/091360  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091478  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091544  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091519  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091626  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091633  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091978  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/091982  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/092182  
PRIOR FILING DATE: 1998-07-09

## Alignment Scores:

Pred. No.:	0.0806	Length:	3265
Score:	91.50	Matches:	46
Percent Similarity:	35.64%	Conservative:	26
Best Local Similarity:	22.77%	Mismatches:	55
Query Match:	8.18%	Indels:	75
DB:	9	Gaps:	8

US-10-034-500-2 (1-218) x US-09-992-598-378 (1-3265)

Qy 17 leuValaApTYrSerGlySerMetMetLyshIsvAlaValaArgGluProLySile 36  
: : : : :  
Db 952 GTTCTGTAAGTCTGGAAGCATGGG----- 978  
Qy 37 GluLeuAlaLySGluAlaIleLeuLySileAenAlaMetProLyMetSerTyGln 56  
: : : : :  
Db 979 -----GGTAAGCAGCCGCTTAATGCAATGCAAGCAGCAAAACATTCTGCTGAG 1032  
Qy 57 GlyLyLeuTyrtThrPheAlaProTySerValIleIleProGlnLySerTyAsnSer 76  
: : : : :  
Db 1033 -----ACTGTGAATAATGATCTGGGTGGG 1059  
Qy 77 CysVal-----AlaGlu 80  
: : : : :  
Db 1060 ATGGTTCACCTTGATAGTACGCCACTATTGTAATAAGCTAATCAATAAAGAGAGT 1119  
Qy 81 CysAlaValaenThrIleLySerAspLeuGluIlePheGlyArgLeuThrProValGly 100  
: : : : :  
Db 1120 GATGAAGAAACACACTCATGCGAGATTACTCATAT-----CCTCTGGGA 1167  
Qy 101 -----AspGlyIleLyMetHisGluThrValIleAenGlnMetProPro 115  
: : : : :  
Db 1168 GGAACCTTCATCTGCTGCGAATTAATATGCAATTCAGGATGGAGAGCTCATTC 1227  
Qy 116 Gln-----AlaAlaValIleLeuLeuThrAspGlyHisAenAsnLeuGlyMetAsn 132



PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090431  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090435  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090444  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090472  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090535  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090540  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090542  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090557  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090676  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090678  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090690  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090694  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090695  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090696  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090862  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/090863  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/091360  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091478  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091544  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091519  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091626  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091633  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091978  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/091982  
PRIOR FILING DATE: 1998-07-07

;; PRIOR APPLICATION NUMBER: 60/092182  
; PRIOR FILING DATE: 1998-07-09  
Alignment Scores:  
Pred. No.: 0.0806 Length: 3265  
Score: 91.50 Matches: 46  
Percent Similarity: 35.64% Conservative: 26  
Best Local Similarity: 22.77% Mismatches: 55  
Query Match: 8.18% Indels: 75  
Gaps: 8  
US-10-034-500-2 (1-218) x US-09-989-293A-378 (1-3265)  
Qy 17 leuValaApTYrSerGlySerMetMetLeuYshisValaValaArgGluProlysile 36  
Db 952 GTTCTGTAAGTCTGGAAGCATGGCG----- 978  
Qy 37 GluLeuAlaLysGluAlaIleLeuLysIleAsnAlaAlaMetProLysMetSerTyrlin 56  
Db 979 -----GGTAAGACCCGCTAAATCGAATGATCAAGCAGCAAAACATTTCCTGCTGCAG 1032  
Qy 57 GlyGlyLeuTyrrhrPheAlaProTyrSerValIleIleProGlnGlySerTyrAsnSer 76  
Db 1033 -----ACTGTGAATGATCCTGGGTGGG 1059  
Qy 77 CysVal-----AlaGlu 80  
Db 1060 ATGGTTCACTTGTAGTACTGCCACTATGTGAATAAGCTAATCAATAAAGCACT 1119  
Qy 81 CysAlaValaAnthrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGly 100  
Db 1120 GATGAAGAAACACACTCATGCGAGGATTAATCAATCTCTGGA 1167  
Qy 101 -----AspGlyIleLysMetHisGluThrValIleAsnGlnMetProPro 115  
Db 1168 GGAACCTTCATCTGCTGCGAATTAATATATGCAATTCAGTGATGGAGACTACATCC 1227  
Qy 116 Gln-----AlaAlaValaIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsn 132  
Db 1228 CAATCGATGATCGAAGTACGATCTGCTGATGCGAGGAGATACACGCAAGTCT 1287  
Qy 133 ProValGluGluValLysSerIleTyrlInThrAsnProAsnValCysPheHisValAl 152  
Db 1288 TGTATTGTAAGTGAA-----CAAGTGGGCCATTGTTCAATTTATTGCTTGG 1338  
Qy 153 SerPheAla----- 155  
Db 1339 GGAAGAGCTGCTGATGAACAGTAATAGAGTGAAGCAAGATACAGAGATCAATTT 1398  
Qy 156 -----AspAspAlaGluGlyLysAlaIleIleAspGlnIleValAlaLeuAsnSer 172  
Db 1399 TATGTTTCAGATGAAGCTCAGAACATGCGCTCATTAATGCTTTGGGGCTTACATCA 1458  
Qy 173 GlySer 174  
Db 1459 GGAAT 1464  
RESULT 7  
US-10-063-547-69  
; Sequence 69, Application US/10063547  
; Publication No. US20020182638A1  
; GENERAL INFORMATION:  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gettitsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; TITLE OF INVENTION: ACIDS ENCODING THE SAME





Oy	17	leuValAlapRyrSerGIysertMetMetCylVhiIvalAlAvalArgLuprolystle	36
Db	952	GTTTGTATTAAGCTGACACCATGGG-----	978
Oy	37	GlueuAlalysglualailleuleylsilaelmalalamelProlysMeSeTyrgln	56
Db	979	-----GGTAGGACCGCCCAATAATCGAAATGAACGACGAAAAATTTCCTGCTGCAG	1032
Oy	57	GlyGlyLeuYrthrPheAlaPProYrservalIIelleProGInglySerThrpsnsEr	76
Db	1033	-----ACTGTGAAAATGATCTCGGGTGGG	1055
Oy	77	CysVal-----Alaclu	80

Db 1060 ATGTTCACTTGTAGTACTGCACTATTGTAAATTAACCTAATCAATAAAGCACT 1119  
Qy 81 CysAlaValAsnThrIleLeuSerAspLeuGluIlePheGlyArgLeuThrProValGly 100  
Db 1120 GATGAAAGAAACACACTCTGCGAGATTACTTACTATAT-----CCTCTGGGA 1167  
Qy 101 -----AspGlyIleLeuSerMetHsgIuThrValIleAsnGlnMetProPro 115  
Db 1168 GGAACCTTCATCTGCTCTGGAATTAATATGATTCATTCAAGTGATTTGAGAGCACTTCC 1227  
Qy 116 Gln-----AlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsn 132  
Db 1228 CAACCTGATGATGATCGAATCGAATCTGCTGCTGATGATGGAGGAGATTAACCTCAAGTTCT 1287  
Qy 123 ProValGluGluValIleLeuSerIleTyrGlnThrAsnProAsnValCysPheHisValVal 152  
Db 1288 TGTATTGATGAAGTGA-----CAAGTGGGGCCATTGTTTATTGCTTTG 1338  
Qy 153 SerPheAla----- 155  
Db 1339 GGAAGAGCTGCTGATGAGCAGATTAATAGATGAGCAAGATTAACAGAGCAAGTCAATTTT 1398  
Qy 156 -----AspAspAlaGluGluValAlaIleIleAspGlnIleValAlaLeuAsnSer 172  
Db 1399 TATGTTTCAAGATGAAGCTGAGAACATGGCCCTTATTGATGCTTTTGGGGCTTTACATCA 1458  
Qy 173 GlySer 174  
Db 1459 GGAAAT 1464

RESULT 9  
US-09-990-444-378  
; Sequence 378, Application US/09990444  
; Publication No. US20020193300A1

## GENERAL INFORMATION:

;; APPLICANT: Ashkenazi, Avi J.  
;; APPLICANT: Baker, Kevin P.  
;; APPLICANT: Botstein, David  
;; APPLICANT: Desnoyers, Luc  
;; APPLICANT: Eaton, Dan L.  
;; APPLICANT: Ferrara, Napoleone  
;; APPLICANT: Fong, Sherman  
;; APPLICANT: Garber, Hanspeter  
;; APPLICANT: Gerltsen, Mary E.  
;; APPLICANT: Goddard, Audrey  
;; APPLICANT: Godowski, Paul J.  
;; APPLICANT: Grimaldi, J. Christopher  
;; APPLICANT: Gurney, Auecin L.  
;; APPLICANT: Kljavin, Ivar J.  
;; APPLICANT: Napier, Mary A.  
;; APPLICANT: Pan, James  
;; APPLICANT: Paoni, Nicholas F.  
;; APPLICANT: Roy, Margaret Ann  
;; APPLICANT: Stewart, Timothy A.  
;; APPLICANT: Tumas, Daniel  
;; APPLICANT: Watanabe, Colin K.  
;; APPLICANT: Williams, P. Mickey  
;; APPLICANT: Wood, William I.  
;; APPLICANT: Zhang, Zemin  
;; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
;; FILE REFERENCE: P2730P1C19  
;; CURRENT APPLICATION NUMBER: US/09/990,444  
;; CURRENT FILING DATE: 2001-11-14  
;; PRIOR APPLICATION NUMBER: 60/049787  
;; PRIOR FILING DATE: 1997-06-16  
;; PRIOR APPLICATION NUMBER: 60/062250  
;; PRIOR FILING DATE: 1997-10-17  
;; PRIOR APPLICATION NUMBER: 60/065186  
;; PRIOR FILING DATE: 1997-11-12  
;; PRIOR APPLICATION NUMBER: 60/065311  
;; PRIOR FILING DATE: 1997-11-13  
;; PRIOR APPLICATION NUMBER: 60/066770

;; PRIOR FILING DATE: 1997-11-24  
;; PRIOR APPLICATION NUMBER: 60/075945  
;; PRIOR FILING DATE: 1998-02-25  
;; PRIOR APPLICATION NUMBER: 60/078910  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/083322  
;; PRIOR FILING DATE: 1998-04-28  
;; PRIOR APPLICATION NUMBER: 60/084600  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/087106  
;; PRIOR FILING DATE: 1998-05-28  
;; PRIOR APPLICATION NUMBER: 60/087607  
;; PRIOR FILING DATE: 1998-06-02  
;; PRIOR APPLICATION NUMBER: 60/087609  
;; PRIOR FILING DATE: 1998-06-02  
;; PRIOR APPLICATION NUMBER: 60/087759  
;; PRIOR FILING DATE: 1998-06-02  
;; PRIOR APPLICATION NUMBER: 60/087827  
;; PRIOR FILING DATE: 1998-06-03  
;; PRIOR APPLICATION NUMBER: 60/088021  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088025  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088026  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088028  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088029  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088030  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088033  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088326  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088167  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088202  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088212  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088217  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088655  
;; PRIOR FILING DATE: 1998-06-09  
;; PRIOR APPLICATION NUMBER: 60/088734  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088738  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088742  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088810  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088824  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088826  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088858  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/088861  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/088876  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/089105  
;; PRIOR FILING DATE: 1998-06-12  
;; PRIOR APPLICATION NUMBER: 60/089440  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089512  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089514  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089532  
;; PRIOR FILING DATE: 1998-06-17

```

; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:
Pred. No.: 0.0806 Length: 3265
Score: 91.50 Matches: 46
Percent Similarity: 35.64% Conservative: 26
Best Local Similarity: 22.77% Mismatches: 55
Query Match: 8.18% Indels: 75
DB: 9 Gaps: 8

US-10-034-500-2 (1-218) x US-09-990-444-378 (1-3265)
QY 17 LeuValAapYrYserGlySerMetMetMetLysHisValAlaValAArgLubProLysIle 36
Db 952 GTTCTGATTAAGTCTGGAAACATCGGG----- 978
QY 37 GluLeuAlaLysGluAlaIleLeuLysIleHisAlaLysMetProLysMetSerTyGln 56
Db 979 -----GGTAAAGACCCCTTAATTCGAGATGAAACAGACGAAACATTTCCTGCTGAG 1032
QY 57 GlyGlyLeuTyThrPheAlaProTySerValIleIleProGlnGlySerThrAsnSer 76
Db 1033 -----ACTGTTGAAAATGAGATCTGGGGGG 1035
QY 77 CysVal-----AlaGlu 80
Db 1060 ATGGTTCATTGATAGTACTGACACTATTGTAAATAGCTAATCAATAAAGACAGT 1119
QY 81 CysAlaValAsnThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGly 100
Db 1120 GATGAAGAAACACACATCCATGAGGATTAACCTACATAT-----CCTCTGGGA 1167
QY 101 -----AspGlyIleLysMetHisGluThrValIleAsnGlnMetProPro 115
Db 1168 GGAACCTTCATCTGCTCTGGAATTTAAATATGACTTTCAAGTGATGAGAGCTACATTCC 1227
QY 116 Gln-----AlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsn 132
Db 1228 CAATCCATGATGATCCGAAGTATCTGCTGCTGATGAGGAGAGATACACTGCCAAGTTCT 1287
QY 133 ProValGluGluValLysSerIleTyGlnThrAsnProAsnValCysPheHisValVal 152
Db 1288 TGTATTTGATGAAGTAA-----CAAAGGGGGCATTTGTTTATTTGCTTTG 1338
QY 153 SerPheAla----- 155
Db 1339 GGAAGAGCTGCTGATGAGACGATTAATAGATGACGACGATTAACAGAGACGATTTT 1398
QY 156 -----AspAspAlaGluGlyLysAlaIleIleAspGlnIleValAlaLeuAsnSer 172
Db 1399 TATGTTTCAGATGAAGCTCAGAAACAATGGCCTCATTTGATGCTTTGGGGCTCTTACATCA 1458
QY 173 GlySer 174
Db 1459 GGAAT 1464

RESULT 10
US-09-989-730-378
; Sequence 378, Application US/09989730
; Publication NO. US20020197674A1
; GENERAL INFORMATION:

```

APPLICANT: Ashkenazi, Avi J.  
 APPLICANT: Baker, Kevin P.  
 APPLICANT: Botstein, David  
 APPLICANT: Desnoyers, Luc  
 APPLICANT: Eaton, Dan L.  
 APPLICANT: Ferrara, Napoleone  
 APPLICANT: Fong, Sherman  
 APPLICANT: Gerber, Hanspeter  
 APPLICANT: Gertlesen, Mary E.  
 APPLICANT: Goddard, Audrey  
 APPLICANT: Godowski, Paul J.  
 APPLICANT: Grimaldi, U. Christopher  
 APPLICANT: Gurney, Austin L.  
 APPLICANT: KJavin, Ivar J.  
 APPLICANT: Napier, Mary A.  
 APPLICANT: Pan, James  
 APPLICANT: Paoni, Nicholas F.  
 APPLICANT: Roy, Margaret Ann  
 APPLICANT: Stewart, Timothy A.  
 APPLICANT: Tumas, Daniel  
 APPLICANT: Watanabe, Colin K.  
 APPLICANT: Williams, P. Mickey  
 APPLICANT: Wood, William I.  
 APPLICANT: Zhang, Zemin  
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 FILE REFERENCE: P27301C69  
 CURRENT FILING DATE: 2001-11-20  
 PRIOR APPLICATION NUMBER: 60/049787  
 PRIOR FILING DATE: 1997-06-16  
 PRIOR APPLICATION NUMBER: 60/062250  
 PRIOR FILING DATE: 1997-10-17  
 PRIOR APPLICATION NUMBER: 60/065186  
 PRIOR FILING DATE: 1997-11-12  
 PRIOR APPLICATION NUMBER: 60/065311  
 PRIOR FILING DATE: 1997-11-13  
 PRIOR APPLICATION NUMBER: 60/066770  
 PRIOR FILING DATE: 1997-11-24  
 PRIOR APPLICATION NUMBER: 60/075945  
 PRIOR FILING DATE: 1998-02-25  
 PRIOR APPLICATION NUMBER: 60/078910  
 PRIOR FILING DATE: 1998-03-20  
 PRIOR APPLICATION NUMBER: 60/083322  
 PRIOR FILING DATE: 1998-04-28  
 PRIOR APPLICATION NUMBER: 60/084600  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/087106  
 PRIOR FILING DATE: 1998-05-28  
 PRIOR APPLICATION NUMBER: 60/087607  
 PRIOR FILING DATE: 1998-06-02  
 PRIOR APPLICATION NUMBER: 60/087609  
 PRIOR FILING DATE: 1998-06-02  
 PRIOR APPLICATION NUMBER: 60/087759  
 PRIOR FILING DATE: 1998-06-02  
 PRIOR APPLICATION NUMBER: 60/087827  
 PRIOR FILING DATE: 1998-06-03  
 PRIOR APPLICATION NUMBER: 60/088021  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088025  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088026  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088028  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088029  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088030  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088033  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088326  
 PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088167  
 PRIOR FILING DATE: 1998-06-05  
 PRIOR APPLICATION NUMBER: 60/088202  
 PRIOR FILING DATE: 1998-06-05  
 PRIOR APPLICATION NUMBER: 60/088212  
 PRIOR FILING DATE: 1998-06-05  
 PRIOR APPLICATION NUMBER: 60/088217  
 PRIOR FILING DATE: 1998-06-05  
 PRIOR APPLICATION NUMBER: 60/088655  
 PRIOR FILING DATE: 1998-06-09  
 PRIOR APPLICATION NUMBER: 60/088734  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088738  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088742  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088810  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088824  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088826  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088858  
 PRIOR FILING DATE: 1998-06-11  
 PRIOR APPLICATION NUMBER: 60/088861  
 PRIOR FILING DATE: 1998-06-11  
 PRIOR APPLICATION NUMBER: 60/088876  
 PRIOR FILING DATE: 1998-06-11  
 PRIOR APPLICATION NUMBER: 60/089105  
 PRIOR FILING DATE: 1998-06-12  
 PRIOR APPLICATION NUMBER: 60/089440  
 PRIOR FILING DATE: 1998-06-16  
 PRIOR APPLICATION NUMBER: 60/089512  
 PRIOR FILING DATE: 1998-06-16  
 PRIOR APPLICATION NUMBER: 60/089514  
 PRIOR FILING DATE: 1998-06-16  
 PRIOR APPLICATION NUMBER: 60/089532  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089538  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089598  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089599  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089600  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089653  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089801  
 PRIOR FILING DATE: 1998-06-18  
 PRIOR APPLICATION NUMBER: 60/089907  
 PRIOR FILING DATE: 1998-06-18  
 PRIOR APPLICATION NUMBER: 60/089908  
 PRIOR FILING DATE: 1998-06-18  
 PRIOR APPLICATION NUMBER: 60/089947  
 PRIOR FILING DATE: 1998-06-19  
 PRIOR APPLICATION NUMBER: 60/089948  
 PRIOR FILING DATE: 1998-06-19  
 PRIOR APPLICATION NUMBER: 60/089952  
 PRIOR FILING DATE: 1998-06-19  
 PRIOR APPLICATION NUMBER: 60/090246  
 PRIOR FILING DATE: 1998-06-22  
 PRIOR APPLICATION NUMBER: 60/090252  
 PRIOR FILING DATE: 1998-06-22  
 PRIOR APPLICATION NUMBER: 60/090254  
 PRIOR FILING DATE: 1998-06-22  
 PRIOR APPLICATION NUMBER: 60/090349  
 PRIOR FILING DATE: 1998-06-23  
 PRIOR APPLICATION NUMBER: 60/090355  
 PRIOR FILING DATE: 1998-06-23  
 PRIOR APPLICATION NUMBER: 60/090429  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090431

```

PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Alignment Scores:
Pred. No.: 0.0806 Length: 3265
Score: 91.50 Matches: 46
Percent Similarity: 35.64% Conservative: 26
Best Local Similarity: 22.77% Mismatches: 55
Query Match: 8.18% Indels: 75
DB: 9 Gaps: 8
US-10-034-500-2 (1-218) x US-09-989-730-378 (1-3265)
QY 17 LeuValAspTyrSerGlySerMetMetCysHisValAlaValArgGluProLysIle 36
Db 952 GTCTTGAATAGCTCTGGAAGCTGGG----- 978
QY 37 GluLeuAlaValGluAlaIleLeuValIleAsnAlaAlaMetProLysMetSerTyrGln 56
Db 979 -----GGTAAGACCGCCTTAATGAATGAATCAAGCAAGCAAAATTTCTGCTGCAG 1032
QY 57 GlyGlyLeuTyrThrPheAlaProTyrSerValIleIleProGlnGlySerTyrPheSer 76
Db 1033 -----ACTGTTGAATAATGATCTCGGTGGG 1059
```

```

QY 77 CysVal-----AlaGlu 80
Db 1060 ATGGTTCACTTGTAGTACTGCGCACTATGTGAAATAGTAATCAATTAAGCAAGT 1119
QY 81 CysAlaValLeuThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGly 100
Db 1120 GATGAAGAAACACACTCATGCGAGGATTAACCTACATAT-----CCTCTGGGA 1167
QY 101 -----AspGlyIleLysMetHisGluThrValIleAsnGlnMetProPro 115
Db 1168 GGAATTCATCTGCTCTGGAATTAATATATGCAATTTACAGGTATTCGAGAGCTACATTC 1227
QY 116 Gln-----AlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsn 132
Db 1228 CAATCGATGATCGCAAGTACTGCTGCTGACTATGAGGAGATTAACATCGCAAGTTCT 1287
QY 133 ProValGluGluValLysSerIleTyrGlnThrAsnProAsnValCysPheHisValVal 152
Db 1288 TGATATGATGAAGTGAAA-----CAAGTGGGCGCATTTTCATTTATGCTTTG 1338
QY 153 SerPheAla----- 155
Db 1339 GGAAGAGCTGCTGATGAAGCACTAATAGAGATGAGCAAGATTAACAGCAAGTATT 1398
QY 156 -----AspAspAlaGluGlyLysAlaIleIleAspGlnIleValAlaLeuAsnSer 172
Db 1399 TATGTTTCAGATGAAGCTCAGAAACATGGCCTCATTTGATGCTTTGGGCTCTTATCA 1458
QY 173 GlySer 174
Db 1459 GGAAT 1464

RESULT 11
US-09-990-436-378
Sequence 378, Application US/09990436
Publication No. US20020198148A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gottlieb, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kijavrin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tuma, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OR INVENTION: Acids Encoding the Same
FILE REFERENCE: P2730P1C14
CURRENT APPLICATION NUMBER: US/09/990,436
PRIOR FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
```



;; PRIOR APPLICATION NUMBER: 60/091478  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091544  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091626  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091633  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091978  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/092182  
;; PRIOR FILING DATE: 1998-07-09

Alignment Scores:  
Pred. No.: 0.0806 Length: 3265  
Score: 91.50 Matches: 46  
Percent Similarity: 35.64% Conservative: 26  
Best Local Similarity: 22.77% Mismatches: 55  
Query Match: 8.18% Indels: 75  
DB: Gaps: 8

US-10-034-500-2 (1-218) x US-09-990-436-378 (1-3265)

Qy 17 LeuValaAprrYrSerGlySerMetMetMetlyShValaAlaValaArgLubProlySile 36  
Db 952 GTTCTGTAAGTCTGGAAGCTGGG----- 978  
Qy 37 GluLeuAlaLysGluAlaIleuLysIleAenAlaAlaMetProlyeweSerTyrGln 56  
Db 979 -----GGAAGGACCGCTTAATGATGATCAAGCAAGAAATTTCTGCTGCG 1032  
Qy 57 GlyGlyLeuYrThrPheAlaProTyrSerValIleIleProGlnGlySerTyrAsnSer 76  
Db 1033 -----ACTGTTGAAATGATGATCCGCGTGGG 1059  
Qy 77 CysVal-----AlaGlu 80  
Db 1060 ATGGTTCCTGTAAGTACGCTCCACTATTGTAATTAAGTAATCAATTAAGCAAGT 1119  
Qy 81 CysAlaValAsnThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGly 100  
Db 1120 GATGAAGAACAACACTATGCGAGATTACTTAATAT-----CTCTGGGA 1167  
Qy 101 -----AspGlyIleLysMetHisGluThrValIleAsnGlnMetProPro 115  
Db 1168 GGAACCTCCATCTGCTGCTGAATTAATATGCAATTCAGTGATGAGAGCTACATTC 1227  
Qy 116 Gln-----AlaAlaValIleLeuLeuThrAspGlyHisAsnGlnLeuGlyMetAsn 132  
Db 1228 CAACCTCGATGATCGGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1287  
Qy 133 ProValGluGluValLysSerIleTyrGlnThrAsnProAsnValCysPheHisValVal 152  
Db 1288 TGTATTGATGAAGTAAA-----CAAGTGGGGCATTTGTTTATTGCTTTG 1338  
Qy 153 SerPheAla----- 155  
Db 1339 GGAAGAGCTGCTGATGAAGCAAGTAATAGATGACCAAGATTAACGAGAGCAATCTTT 1398  
Qy 156 -----AspAspAlaGluGluGlyLysAlaIleIleAspGlnIleValAlaLeuAsnSer 172  
Db 1399 TATGTTTCAATGAAGCTCAGAACATGCGCTCATTTGCTTTGGGCTCTTACATCA 1458  
Qy 173 GlySer 174  
Db 1459 GGAAT 1464  
RESULT 12  
US-09-991-181-378

Sequence 378, Application US/09991181  
Publication No. US20020197615A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi J.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Deenoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gottfries, Mary B.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Collin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OR INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2730PIC53  
CURRENT FILING DATE: 2001-11-16  
PRIOR APPLICATION NUMBER: US/09/991,181  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 2001-11-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033



PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23

PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090431  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090435  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090444  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090472  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090535  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090540  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090542  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090557  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090676  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090678  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090690  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090694  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090695  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090696  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090862  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/090863  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/091360  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091478  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091544  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091519  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091626  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091633  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091978  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/091982  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/092182  
PRIOR FILING DATE: 1998-07-09

## Alignment Scores:

Pred. No.:	0.0806	Length:	3265
Score:	91.50	Matches:	46
Percent Similarity:	35.64%	Conservative:	26
Best Local Similarity:	22.77%	Mismatches:	55
Query Match:	8.18%	Indels:	75
DB:	9	Gaps:	8

US-10-034-500-2 (1-218) x US-09-991-181-378 (1-3265)

Qy 17 LeuValAspTySerGlySerMetMetLySHsValAlaValArgGluProLysIle 36  
Db 952 GTTCTGATAGCTCGAGCATGGG----- 978  
Qy 37 GluLeuAlaGluAlaIleLeuYsIleAsnAlaIleMetProLysMetSerTyrgIn 56  
Db 979 -----GTTAAGACCGCCTTAATGAAATCAAGACGAAACATTTCCTGCTGCG 1032

Qy 57 GlyGlyLeuYrThrPheAlaProTyrSerValIleIleProGlnGlySerTrpAsnSer 76  
Db 1033 -----ACTGTTGAAATGATCTCGGTGGG 1059  
Qy 77 CysVal -----AlaGlu 80  
Db 1060 ATGGTTCCTTGGATAGTACTGCCACTATTGTAATTAAGCTAATCAATAAAGCACT 1119  
Qy 81 CysAlaValAsnThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGly 100  
Db 1120 GATGAAGAACAACACTCATGCGAGATTACTACATAT-----CCTCTGGGA 1167  
Qy 101 -----AspGlyIleLysMetHisGluThrValIleAsnGlnMetProPro 115  
Db 1168 GCAACTTCCTCATCTGCTCTGCAATTAAATATGCAATTTCAGTGATTGAGAGCTCATTTCC 1227  
Qy 116 Gln-----AlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsn 132  
Db 1228 CAAGCTCGATGCATCCGAAGTACTGCTGCTGACTGATGGGAGATACACTGCAAGTTCT 1287  
Qy 133 ProValGluGluValLysSerIleTyrGlnThrAsnProAsnValCysPheHisValVal 152  
Db 1288 TGTATTGATGAAGTAA-----CAAGTGGGGCATTGTTTATTTCCTTG 1338  
Qy 153 SerPheAla----- 155  
Db 1339 GCAAGAGCTGCTGATGAAGCACTAATAGATAGACCAAGATTAACAGAGCAAGCTATTT 1398  
Qy 156 -----AspAspAlaGluGlyLysAlaIleIleAspGlnIleValAlaLeuAsnSer 172  
Db 1399 TATGTTTCAATGAAAGCTCAGAACTAGGAGCTCATTTGAGGCTCTTACATCA 1458  
Qy 173 GlySer 174  
Db 1459 GGAAT 1464

RESULT 13  
US-09-993-687-378  
Sequence 378, Application US/09993687  
Publication No. US20020198149A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi J.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Deenoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerltsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Guiney, Auectin L.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zenlin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2730P1C11  
CURRENT APPLICATION NUMBER: US/09/993,687  
PRIOR FILING DATE: 2002-11-14  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250

PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16

PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090431  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090435  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090444  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090472  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090535  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090540  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090542  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090578  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090650  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090694  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090695  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090696  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090862  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/090863

PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/091360  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091478  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091544  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091519  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091626  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091633  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091978  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/091982  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/092182  
PRIOR FILING DATE: 1998-07-09

## Alignment Scores:

Pred. No.:	0.0806	Length:	3265
Score:	91.50	Matches:	46
Percent Similarity:	35.64%	Conservative:	26
Best Local Similarity:	22.77%	Mismatches:	55
Query Match:	8.18%	Indels:	75
DB:	9	Gaps:	8

US-10-034-500-2 (1-218) x US-09-993-687-378 (1-3265)

QY 17 LeuValaapTySerGlySerMetMetLysHisValaAlaValArgGluProLysIle 36  
Db 992 GTTCTGTAATAGTCTGCAAGCATGGCG----- 978  
QY 37 GluLeuAlaIysGluAlaIleLeuLysIleLeuAlaAlaMetProLysMetSerTyGln 56  
Db 979 -----GGTAAGACCGCTCAATCGAATGCAAGCAAAACATTCTCTGCTGAG 1032  
QY 57 GlyGlyLeuTyThrPheAlaProTySerValIleIleProGlnGlySerTyPasnSer 76  
Db 1033 -----ACTGTGAAATGATCTCTGGGTGGG 1059  
QY 77 CysValI-----AlaGlu 80  
Db 1060 ATGTTCACTTGTATAGTACGCACTATTGTAATAGCTAATCAATAAAGCAGT 1119  
QY 81 CysAlaValaSerThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGly 100  
Db 1120 GATGAAGAAACACACTGATGCGAGATTACCTACATAT-----CTCTGGGA 1167  
QY 101 -----AspGlyIleLysMetHisGluThrValIleAsnGlnMetProPro 115  
Db 1168 GGAACCTTCATGCTGCTCGAATTAAATATGATCATTTGAGGTATGGAGACTACATTCC 1227  
QY 116 Gln-----AlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsn 132  
Db 1228 CAACGATGATGATCCGAAGTACTGCTGCTGATGATGGGAGATAAACATGCAATTCT 1287  
QY 133 ProValGluGluValIlySerIleTyGlnThrAsnProAsnValCysPheHisValVal 152  
Db 1288 TGTATTGATGAAGTAAAA-----CAAGTGGGGCCATTGTTATTTGCTTTG 1338  
QY 153 SerPheAla----- 155  
Db 1339 GGAAGAGTCTGATGAGCACTAATAGATGACCAAGATACAGAGGAAGTCAATTT 1398  
QY 156 -----AspAspAlaGluGlyLysAlaIleIleAspGlnIleValAlaLeuAsnSer 172  
Db 1399 TATGTTTCAGATGAAGCTCAGAAACATGGCCCTCATTTGATGCTTTGGGCTCTTACATCA 1458  
QY 173 GlySer 174  
Db 1459 GGAAT 1464

RESULT 14  
US-09-989-734-378  
Sequence 378, Application US/09989734  
Publication No. US20030003531A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi J.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Deanoeyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerltsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Goddard, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gutney, Austin L.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Wacanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2730P1C64  
CURRENT APPLICATION NUMBER: US/09/989,734  
PRIOR FILING DATE: 2001-11-19  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349

PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090431  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090435  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090444  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090472  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090535  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090540  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090542  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090557  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090676  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090678  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090690  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090694  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090695  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090696  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090862  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/090863  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/091360  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091478  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091444  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091519  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091626  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091633  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091978  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/091982  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/092182  
PRIOR FILING DATE: 1998-07-09

## Alignment Scores:

Pred. No.: 0.0806 Length: 3265  
Score: 91.50 Matches: 46  
Percent Similarity: 35.64 Conservative: 26  
Best Local Similarity: 22.77 Mismatches: 55  
Query Match: 8.18 Indels: 75  
DB: 9 Gaps: 8

US-10-034-500-2 (1-218) x US-09-989-734-378 (1-3265)

QY 17 LeuValAspTyrSerGlySerMetMetMetLeuHisValAlaValAlaArgGluProLysIle 36  
DB 952 GTTCTGATTAAGTCTGGAACATGGGG----- 978  
QY 37 GluLeuAlaLysGluAlaIleLeuLysIleAsnAlaLametProLysMetSerTyrGln 56

DB 979 -----GTTAGGACCGCTTAATCGAATGATCAACAGCAAAATTCCTGCTGACG 1032  
QY 57 GlyGlyLeuTyrThrPheAlaProTyrSerValIleIleProGlnGlySerTyrPanser 76  
DB 1033 -----ACTGTTGAAATGATCTGGGTGGG 1059  
QY 77 CysVal-----AlaGlu 80  
DB 1060 ATGTTTCACTTGTGATGACTGCACCTATTGTAAATAAGCTAATCAATTAAGAGCT 1119  
QY 81 CysAlaValAsnThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGly 100  
DB 1120 GATGAAGAAACACACTCATGCTGACGAGATTACCTACATAT-----CCTGTGGGA 1167  
QY 101 -----AspGlyIleLysMetHisGluThrValIleAsnGlnMetProPro 115  
DB 1168 GGAATTCATCTGCTGCTGATTAATATGCAATTCAGGGATTTGAGAGCTACATTC 1227  
QY 116 Gln-----AlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsn 132  
DB 1228 CACTCGATGATCGGAAGTACTGCTGCTGACTGATGGGAGATTAACACTGCAAGTTCT 1287  
QY 133 ProValGluGluValLysSerIleTyrGlnThrAsnProAsnValCysPheHisValVal 152  
DB 1288 TGTATTGATGAAGTGA-----CAAGTGGGGCCATTGTCATTTATTGCTTTG 1338  
QY 153 SerPheAla----- 155  
DB 1339 GGAGAGCTGCTGATGAACGATTAATGAGATGACGAATTAACAGGAGAGATCTTT 1398  
QY 156 -----AspAspAlaGluGlyLysAlaIleIleAspGlnIleValAlaLeuAsnSer 172  
DB 1399 TATGTTGATGAAGCTCAGAACATGCCCTGATGCTTTGGGGCTCTTACATCA 1458  
QY 173 GlySer 174  
DB 1459 GGAAT 1464  
RESULT 15  
US-09-997-653-378  
; Sequence 378, Application US/09997653  
; Publication No. US2003008297A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kijavini, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tunas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2730PTC38  
CURRENT APPLICATION NUMBER: US/09/997,653  
CURRENT FILING DATE: 2001-11-15

[illegible]

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

Alignment Scores:	
Pred. No.:	0.0806
Score:	91.50
Percent Similarity:	35.64%
Best Local Similarity:	22.77%
Query Match:	8.18%
DB:	9
Length:	3266
Matches:	46
Conservative:	26
Mismatches:	55
Indels:	75
Gaps:	8

US-10-034-500-2 (1-218) X US-09-997-653-378 (1-3265)

Oy	17	LeuValAspArgySerGlySerMetMetMetLeuHisIleValAlaValAlaGlnProLeu	36
Db	952	GTCTCTGTAATAGCTCTGGAGGATGGG-----	978
Oy	37	GluLeuValAlaLysGluAlaIleLeuLysIleAsnAlaIleMetProLysMetSerTyGln	56
Db	979	-----GGTAAGGAGCCGCCCTTAATCGAATGAATCAAGACGAAACAAATTTCTCGTGGAG	10322
Oy	57	GlyGlyLeuTyThrPheAlaProTySerValIleIleProGlnGlySerTrpAsnSer	76
Db	1033	-----ACTGTTGAATAATGAGATCTGGGCTGGGAG	10595
Oy	77	CysVal-----AlaGln	80
Db	1060	ATGGTCACTTGTAGTACTGCCACTATTTGTAAATAGCTAAATCCAAATAAAGGAGCT	1119
Oy	81	CysAlaValAsnThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGly	100
Db	1120	GATGAAGAAGAACACACTCATGCGAGATTTCCATACATAT-----CCTCTGGGA	1167
Oy	101	-----AspGlyIleLysMetHisGluThrValIleAsnGlnMetProPro	115
Db	1168	GGAACTTCATCTGCCTCGGAAATTAATATGCAATTTCCAGCTGTATGGAGAGCTACATATCC	1227
Oy	116	Gln-----AlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsn	132
Db	1228	CAACTCGATGATCGAGAGATACCTCTCTGCTGATGAGGAGATACACCTGCMAAGTTT	1287
Oy	133	ProValGluGluValLysSerIleTyGlnThrAsnProAsnValCysPheHisValVal	152
Db	1288	TGTATGTAGTAGAAGTAA-----CAAGTGGGGCCATTTGCTATTTTATTTGGCTTTG	1338
Oy	153	SerPheAla-----	155
Db	1339	GGAGAGCTGCTGATGAACAGTAAATAGATGAGCAAGATTAACAGAGAAATCATTTT	1399
Oy	156	-----AspAspAlaGluGlyLysAlaIleIleLeuAspGlnIleValAlaLeuAsnSer	172
Db	1399	TATGTTTCAGATGAGCTTCAGAACATAGGCTCATATGATCTTTTGGGGCTTTACATCA	1456

QY	173	GLYSer	174
		:::	
Db	1459	GGAAAT	1464

RESULT 16  
US-10-174-590-257

```

; Sequence 257, Application US/10174590
; Publication No. US20030008352A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Deenoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zhenxi
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; ACIDS ENCODING THE SAME
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C42
; CURRENT APPLICATION NUMBER: US/10/174,590
; CURRENT FILING DATE: 2002-06-18
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 257
; LENGTH: 3265
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-174-590-257

```

US-10-034-500-2 (1-218) X US-10-174-590-257 (1-3265)

```

QY 17 LeuValAspTrySerGlySerMetMetMetuLySHLeValAlaValArgGluProIysIle 36
Db 952 GTTCTGTTAAAGTCTGGAGCATGGGG----- 978
QY 37 GluLeuAlaIysGluAlaIleLeuLysIleAsnAlaAlaMetProIysMetSerTyGln 56
Db 979 -----GGTAAGAGACCCCTTAATCAATGAATGAACAGCAAGAAAATTCTCGTGTGAG 1032
QY 57 GlyGlyLeuTyThrPheAlaProTySerValIleIleProGlnGlySerTrpAsnSer 76
Db 1033 -----ACTGTTGAAAATGATCTGGGTGGG 1059
QY 77 CysVal-----AlaGlu 80
Db 1060 ATGGTTCACITTTGATAGTACTGCCACTATTGTAATAGCTAATCAATAAAGAGCGGT 1119
QY 81 CysAlaValAsnThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGly 100
Db 1120 GATGAAGAAACACACTCAGTCAGCGAGATTTCCTACATAT-----CCTTGGGA 1167
QY 101 -----AspGlyIleLysMetHisGluThrValIleAsnGlnMetProPro 115
Db 1168 GGAACCTTCATCTGCTCTGGAGATTAATATGACTTCAGGTGATTGGAGACTACATTCC 1227
QY 116 Gln-----AlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsn 132
Db 1228 CAATCGATGATGCCGAATGACTCTGCTGACTGATGAGGAGGATTAACCTCGCAAGTTCT 1287
QY 133 ProValGluGluValIysSerIleTyrglnThrAsnProAsnValCysPheHisValVal 152

```

Db 1288 TGTATTGATGAAGTGA--CAAGTGGGCCATTGTTTATTATGCTTTG 1338  
 Qy 153 Serphea-----155  
 Db 1339 GGAAGAGTGTGATGAGCACTAATAGATGACGAAATACAGAGAACTATTTT 1398  
 Qy 156 -----AaPaPaPAlaGlugllyLysAlaIlelleaPpGlnlleValAlaLeuAnsSer 172  
 Db 1399 TATGTTTCAGATGAAGCTCAGAACAAATGCGCTCATTTGCTTTGGGGCTCTTACATCA 1458  
 Qy 173 GlySer 174  
 Db 1459 GGAAT 1464

RESULT 17  
 US-10-176-758-257  
 ; Sequence 257, Application US/10176758  
 ; Publication No. US2003008353A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Chen, Jian  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Pan, James  
 ; APPLICANT: Smith, Victoria  
 ; APPLICANT: Watanabe, Colin K.  
 ; APPLICANT: Wood, William I.  
 ; APPLICANT: Zhang, Zemin  
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ; FILE REFERENCE: P3430R1C104  
 ; CURRENT APPLICATION NUMBER: US/10/176,758  
 ; CURRENT FILING DATE: 2002-06-21  
 ; Prior Application removed - See File Wrapper or Palm  
 ; NUMBER OF SEQ ID NOS: 612  
 ; SEQ ID NO 257  
 ; LENGTH: 3265  
 ; TYPE: DNA  
 ; ORGANISM: Homo Sapien  
 US-10-176-758-257

Alignment Scores:  
 Pred. No.: 0.0806 Length: 3265  
 Score: 91.50 Matches: 46  
 Percent Similarity: 35.64% Conservative: 26  
 Best Local Similarity: 22.77% Mismatches: 55  
 Query Match: 8.18% Indels: 75  
 Gaps: 8

US-10-034-500-2 (1-218) x US-10-176-758-257 (1-3265)  
 Qy 17 LeuValaApTYrSerGlySerMetMetClyshisValAlaValaArgLubProlysile 36  
 Db 952 GTTCTTGAATAGCTGGAAGCATGGG-----978  
 Qy 37 GluLeuAlaLysGluAlaIleLeuLysIleAenAlaAlaMetProlyMetSerTyrgln 56  
 Db 979 -----GGTAAGACCGCCTTAATGAATCAAGCAAAACATTTCCTGCTGCAG 1032  
 Qy 57 GlyGlyLeuTyrrThpheaLapProTySerValIlelleProGlnGlySerTyraAnsSer 76  
 Db 1033 -----ACTGTTGAAAATGATCCTCGGTGGGG 1059  
 Qy 77 CysVal-----AlaGlu 80  
 Db 1060 ATGTTCACTTGTATAGTACCTGCACTATTGTAAATAGCTAATCAATAAAGCAGT 1119  
 Qy 81 CysAlaValaenThrlleLysSerSerpLeuGluIlePheGlyArgLeuThrProValGly 100  
 Db 1120 GATGAAGAAACACACTATGCGAGGATTAACCTACATAT-----CCTCTGGGA 1167

Qy 101 -----AaPpGlylleLysMetClyshThrValIleAenGlnMetProPro 115  
 Db 1168 GAACTTCATCTGCTCTCGAATTAATATGCAATTCAGGTAGTGAAGCTACATTC 1227  
 Qy 116 Gln-----AlaAlaValIleLeuLeuThrAaPpGlyshsAaenLeuGlyMetAsn 132  
 Db 1228 CAATTCATGATGATGAGATGACTGCTGCTGATGAGGAGATPAACATGCAATTCCT 1287  
 Qy 133 ProValGluGluValLysSerIleTyrglnThraenProAenValCysPhehIleVal 152  
 Db 1288 TGTATTGATGAAGTGA--CAAGTGGGCCATTGTTTATTATGCTTTG 1338  
 Qy 153 Serphea-----155  
 Db 1339 GGAAGAGTGTGATGAGCACTAATAGATGACGAAATACAGAGAACTATTTT 1398  
 Qy 156 -----AaPaPaPAlaGlugllyLysAlaIlelleaPpGlnlleValAlaLeuAnsSer 172  
 Db 1399 TATGTTTCAGATGAAGCTCAGAACAAATGCGCTCATTTGCTTTGGGGCTCTTACATCA 1458  
 Qy 173 GlySer 174  
 Db 1459 GGAAT 1464

RESULT 18  
 US-10-063-616-69  
 ; Sequence 69, Application US/10063616  
 ; Publication No. US20030013855A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Eaton, Dan L.  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Gerlitsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Grimaldi, Christopher J.  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Watanabe, Colin K.  
 ; APPLICANT: Wood, William I.  
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ; FILE REFERENCE: P3230R1C1  
 ; CURRENT APPLICATION NUMBER: US/10/063,616  
 ; CURRENT FILING DATE: 2002-05-03  
 ; Prior Application removed - See File Wrapper or Palm  
 ; NUMBER OF SEQ ID NOS: 170  
 ; SEQ ID NO 69  
 ; LENGTH: 3265  
 ; TYPE: DNA  
 ; ORGANISM: Homo Sapien  
 US-10-063-616-69

Alignment Scores:  
 Pred. No.: 0.0806 Length: 3265  
 Score: 91.50 Matches: 46  
 Percent Similarity: 35.64% Conservative: 26  
 Best Local Similarity: 22.77% Mismatches: 55  
 Query Match: 8.18% Indels: 75  
 Gaps: 8

US-10-034-500-2 (1-218) x US-10-063-616-69 (1-3265)  
 Qy 17 LeuValaApTYrSerGlySerMetMetClyshisValAlaValaArgLubProlysile 36  
 Db 952 GTTCTTGAATAGCTGGAAGCATGGG-----978  
 Qy 37 GluLeuAlaLysGluAlaIleLeuLysIleAenAlaAlaMetProlyMetSerTyrgln 56  
 Db 979 -----GGTAAGACCGCCTTAATGAATCAAGCAAAACATTTCCTGCTGCAG 1032  
 Qy 57 GlyGlyLeuTyrrThpheaLapProTySerValIlelleProGlnGlySerTyraAnsSer 76  
 Db 1033 -----ACTGTTGAAAATGATCCTCGGTGGGG 1059



Qy	77	CysVal	-----	AlaIleu	80
				:::	
Db	1060	ATGGTACTCTTGATGTAAGTACTGCCACTATTGTAAATTAACCTTAATCCAAATAAAGACAGT			1119
Qy	81	CysAlaValAsnThrIleLeuSerAspLeuGluIlePheGlyArgLeuThrProValGly			100
			:::		
Db	1120	GATGAAGAAGAAACAACATCCATCGAGGAGTATCCATACATAT	-----	CCCTCGGGA	1167
Qy	101	-----	ArgGlyIleLeuSerMetHisGluThrValIleAsnGlnMetProPro		115
Db	1168	GGAACCTTCATCTGCTCGCTCGGAATTAATATATGATTCACAGTGTATGGAGAAGCTACATTC			1227
Qy	116	Gln-----	-AlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsn		132
		:::			:::
Db	1228	CAATCGATGATATCCGAAGATACGTCTGCTGATCGATCGGGAGATATACCTGCAAGTTCT			1287
Qy	133	ProValGluGluValIlySerIleTyrGlnThrAsnProAsnValCysPheHisValVal			152
					:::
Db	1288	TGTATGTATGAAAGTGAAT-----	CAAGTGGGGCCATGTTTCATTTTATTTGCTTTG		1338
Qy	153	SerPheIle	-----		155
Db	1339	GGAAAGCTGCTGATGAGCACTAATATAGAGTAGACACAGTAAACAGAGGAAGTCAATTTT			1398
Qy	156	-----	-AspAspAlaGluGlyIlySalaIleIleAspGlnIleValAlaLeuAsnSer		172
Db	1399	TATGTTTCAGATGAAAGCTCAGAACAAATGGCTCATTTGATGCTTTTGGGCTCTTACATCA			1458
Qy	173	GlySer	174		
Db	1459	GGAAAT	1464		

[illegible]

PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090696  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090862  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/090863  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/091360  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091478  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091544  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091519  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091626  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091633  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091978  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/091982  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/092182  
PRIOR FILING DATE: 1998-07-09

## Alignment Scores:

Pred. No.:	0.0806	Length:	3265
Score:	91.50	Matches:	46
Percent Similarity:	35.64%	Conservative:	26
Best Local Similarity:	22.77%	Mismatches:	55
Query Match:	9	Indels:	75
		Gaps:	8

US-10-034-500-2 (1-218) x US-09-993-667-378 (1-3265)

Qy 17 LeuValAspTyrSerGlySerMetMetLysHisValAlaValArgLupProLysIle 36  
Db 952 GTTCTTGATAGTCTGGAAGCATGGG----- 978  
Qy 37 GluLeuAlaLysGluAlaIleLeuLysIleAsnAlaIleMetProLysMetSerTyrgln 56  
Db 979 -----GGTAAGACCGCGCTTAATCGAATGCAAGCAAAACATTTCCTGCTGCAG 1032  
Qy 57 GlyGlyLeuTyThrPheAlaProTyrservAlaIleIleProGlnGlySerTrpAsnser 76  
Db 1033 -----ACTGTTGAAAATGATCTCGGTGGG 1059  
Qy 77 CyVal-----AlaGln 80  
Db 1060 ATGGTTCACTTGATAGTACTGCGCACTATTGTAATAGCTAATCCAAATAAAGCAGT 1119  
Qy 81 CyAlaValAsnThrIleLysSerAspLeuGlnIlePheGlyArgLeuThrProValGly 100  
Db 1120 GATGAAGAAGAACACACTGTCAGCAGATTACCTACATAT-----CCTCGGGA 1167  
Qy 101 -----AspGlyIleLysMetHisGluThrValIleAsnGlnMetProPro 115  
Db 1168 GGAATTCATCTGCTCTGGAATTAATATGCAATTTCAGGTGATGAGAGCTACATTC 1227  
Qy 116 Gln-----AlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsn 132  
Db 1228 CAACGTGATGATCGCAAGTACTGCTGCTGACTGATGAGGAGATTAACCTGCAAGTCT 1287  
Qy 133 ProValGluGluValLysSerIleTyrglnThrAsnProAsnValCySPheHisValVal 152  
Db 1288 TGAATTGATGAGTGAATA-----CAAAATGGGGCATGTGTCACTTTATTGCTTTG 1338  
Qy 153 SerPheAla----- 155  
Db 1339 GGAAGAGCTGCTGATGAAGCATTAAGATGAGCAAGATTAACAGGAGAGCATTTT 1398  
Qy 156 -----AspAspAlaGluGlyLysAlaIleIleAspGlnIleValAlaLeuAsnser 172

Db 1399 TATGTTTCAGATGAAGTCAAGAACATGCGCTCATGTGCTTTGGGCTCTTACATCA 1458  
Qy 173 GlySer 174  
Db 1459 GGAAT 1464

## RESULT 21

US-10-063-502-69  
Sequence 69, Application US/10063502  
Publication No. US20030025042A1  
GENERAL INFORMATION:  
APPLICANT: Eaton, Dan L.  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gerlitsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, Christopher J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Macanabe, Colin K.  
APPLICANT: Wood, William I.  
TITLE OR INVENTION: ACIDS ENCODING THE SAME  
FILE REFERENCE: P3230R1C1  
CURRENT FILING DATE: 2002-05-01  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 170  
SEQ ID NO 69  
LENGTH: 3265  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-063-502-69

## Alignment Scores:

Pred. No.:	0.0806	Length:	3265
Score:	91.50	Matches:	46
Percent Similarity:	35.64%	Conservative:	26
Best Local Similarity:	22.77%	Mismatches:	55
Query Match:	9	Indels:	75
		Gaps:	8

US-10-034-500-2 (1-218) x US-10-063-502-69 (1-3265)

Qy 17 LeuValAspTyrSerGlySerMetMetLysHisValAlaValArgLupProLysIle 36  
Db 952 GTTCTTGATAGTCTGGAAGCATGGG----- 978  
Qy 37 GluLeuAlaLysGluAlaIleLeuLysIleAsnAlaIleMetProLysMetSerTyrgln 56  
Db 979 -----GGTAAGACCGCGCTTAATCGAATGCAAGCAAAACATTTCCTGCTGCAG 1032  
Qy 57 GlyGlyLeuTyThrPheAlaProTyrservAlaIleIleProGlnGlySerTrpAsnser 76  
Db 1033 -----ACTGTTGAAAATGATCTCGGTGGG 1059  
Qy 77 CyVal-----AlaGln 80  
Db 1060 ATGGTTCACTTGATAGTACTGCGCACTATTGTAATAGCTAATCCAAATAAAGCAGT 1119  
Qy 81 CyAlaValAsnThrIleLysSerAspLeuGlnIlePheGlyArgLeuThrProValGly 100  
Db 1120 GATGAAGAAGAACACACTGTCAGCAGATTACCTACATAT-----CCTCGGGA 1167  
Qy 101 -----AspGlyIleLysMetHisGluThrValIleAsnGlnMetProPro 115  
Db 1168 GGAATTCATCTGCTCTGGAATTAATATGCAATTTCAGGTGATGAGAGCTACATTC 1227  
Qy 116 Gln-----AlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsn 132  
Db 1228 CAACGTGATGATCGCAAGTACTGCTGCTGACTGATGAGGAGATTAACCTGCAAGTCT 1287  
Qy 133 ProValGluGluValLysSerIleTyrglnThrAsnProAsnValCySPheHisValVal 152

```

Db 1288 TGTATTGTAAGTAAAA-----CAAGTGGGGCCATTGTCATTATTATGCTTTG 1338
Qy 153 Serphea1----- 155
Db 1339 GGAAGAGCTGCTGATGAAGCAGTAATAGAGATGACAGATAACAGAGAACTATT 1398
Qy 156 -----AaPaPaLaGlUgLyLaLaLeLeLeaPpGlnLeVaLaLeuAnSer 172
Db 1399 TATGTTTCAGATGAAGCTCAGAACATGCGCTCATGATGCTTTGGGCTCTTACATCA 1458
Qy 173 GlySer 174
Db 1459 GGAAT 1464

RESULT 22
US-10-173-706-257
; Sequence 257, Application US/10173706
; Publication No. US20030022293A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Deanovers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C7
; CURRENT APPLICATION NUMBER: US/10/173,706
; PRIOR FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 257
; LENGTH: 3265
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-173-706-257

Alignment Scores:
Pred. No.: 0.0806 Length: 3265
Score: 91.50 Matches: 46
Percent Similarity: 35.64% Conservative: 26
Best Local Similarity: 22.77% Mismatches: 55
Query Match: 8.18% Indels: 75
DB: Gaps: 8

US-10-034-500-2 (1-218) x US-10-173-706-257 (1-3265)
Qy 17 LeuValaApTySerGlySerMetMetMetLyshIValaLaValaArgGluProLysIle 36
Db 952 GTTCTTGATTAAGCTCGAAGCATGGCG----- 978
Qy 37 GluLeuAlaLySGluAlaLeLeuLySleLeuAlaLaMetProLysMetSerTyrgln 56
Db 979 -----GSTAAAGACCGCTCAATATGAAATGCAAGCAAAACATTTCTGCTGCAG 1032
Qy 57 GlyGlyLeuTyRThrPheAlaProTySerValIleIleProGlnGlySerTyraPanser 76
Db 1033 -----ACTGTTGAAAATGAGATCTGGGTGGGG 1059
Qy 77 CysVal-----AlaGlu 80
Db 1060 ATGTTCACTTGATAGTACGCACTATTGTAAATTAAGCTAATCCAATMAAAGCAGT 1119
Qy 81 CysAlaValaPheThrIleLeuSerSerLeuGlnIlePheGlyArgLeuThrProValGly 100.
Db 1120 GATGAAGAAACACACTCATGCGAGATTACCTACATAT-----CCTCTGGGA 1167

```

```

Qy 101 -----AaPpLyIleLySmetChIsgIurThValIleAnGlnMetProPro 115
Db 1168 GGAACCTCCATCTGCTCGAATTAATATGACATTTACAGTGATTCGAGAGCTACATCC 1227
Qy 116 Gln-----AlaAlaValIleLeuLeuThrAaPpGlyhIshAnSerLeuGlyMetAn 132
Db 1228 CAACTCGATGATCGGAATGCTGCTGCTGATGATGGGAGGATTAACATCGCAAGTTCT 1287
Qy 133 ProValGluGluValLySerIleTyrglnThrAnPheAnValCysPhehIleVala 152
Db 1288 TGTATTGTAAGTAAAA-----CAAGTGGGGCCATTGTCATTATTATGCTTTG 1338
Qy 153 Serphea1----- 155
Db 1339 GGAAGAGCTGCTGATGAAGCAGTAATAGAGATGACAGATAACAGAGAACTATT 1398
Qy 156 -----AaPaPaLaGlUgLyLaLaLeLeLeaPpGlnLeVaLaLeuAnSer 172
Db 1399 TATGTTTCAGATGAAGCTCAGAACATGCGCTCATGATGCTTTGGGCTCTTACATCA 1458
Qy 173 GlySer 174
Db 1459 GGAAT 1464

RESULT 23
US-10-175-738-257
; Sequence 257, Application US/10175738
; Publication No. US20030022294A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Deanovers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C45
; CURRENT APPLICATION NUMBER: US/10/175,738
; PRIOR FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 257
; LENGTH: 3265
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-175-738-257

Alignment Scores:
Pred. No.: 0.0806 Length: 3265
Score: 91.50 Matches: 46
Percent Similarity: 35.64% Conservative: 26
Best Local Similarity: 22.77% Mismatches: 55
Query Match: 8.18% Indels: 75
DB: Gaps: 8

US-10-034-500-2 (1-218) x US-10-175-738-257 (1-3265)
Qy 17 LeuValaApTySerGlySerMetMetMetLyshIValaLaValaArgGluProLysIle 36
Db 952 GTTCTTGATTAAGCTCGAAGCATGGCG----- 978
Qy 37 GluLeuAlaLySGluAlaLeLeuLySleLeuAlaLaMetProLysMetSerTyrgln 56
Db 979 -----GSTAAAGACCGCTCAATATGAAATGCAAGCAAAACATTTCTGCTGCAG 1032
Qy 57 GlyGlyLeuTyRThrPheAlaProTySerValIleIleProGlnGlySerTyraPanser 76

```

```

Db 1033 -----ACTGTTGAAAATGAAATCTCGGTGGGG 1059
Oy 77 CysVal-----AlaGlu 80
Db 1060 ATGGTTCACTTGTATGACTGCCACTATTGTAAATAGCTAATCAATAAAAGCAGT 1119
Oy 81 CysAlaValAanThrIleLeuSerAspLeuGluIlePheGlyArgLeuThrProValGly 1100
Db 1120 GATGAAAGAAACACACTGTCGACAGATTACCTACATAT-----CCTCTGGGA 1167
Oy 101 -----AspGlyIleLeuSerHisGluThrValIleAsnGlnMetProPro 115
Db 1168 GGAACCTTCATCTGCTCGTGAATTAATATGCAATTCAGGTGATTGAGAGCACTACATTC 1227
Oy 116 Gln-----AlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsn 132
Db 1228 CAACTCGATGATCCGAACTGCTGCTGACTGATGGGAGATAAACCTGCAAGTTCT 1287
Oy 133 ProValGluGluValIleYserIleTyrglnThrAsnProAsnValCysPheHisValVal 152
Db 1288 TGTATTGATGAAGTGAAA-----CNAAGTGGGGCCATTGTTTATTGCTTTG 1338
Oy 153 SerPheAla----- 155
Db 1339 GGAAGAGCTGCTGATGAACAGTAATAGATGAGCAAGATAACAGAGAGTCAATTTT 1398
Oy 156 -----AspAspAlaGluGlyLysAlaIleIleAspGlnIleValAlaLeuAsnSer 172
Db 1399 TATGTTTCAGATGAAGCTCAGAACATGGCCTCATTTGATGCTTTGGGGCTCTTACATCA 1458
Oy 173 GlySer 174
Db 1459 GGAAT 1464

```

# RESULT 24

US-10-175-752-257

; Sequence 257, Application US/10175752

; Publication No. US20030022295A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430R1C60

; CURRENT APPLICATION NUMBER: US/10/175,752

; CURRENT FILING DATE: 2002-06-19

; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 257

; LENGTH: 3265

; TYPE: DNA

; ORGANISM: Homo Sapien

US-10-175-752-257

## Alignment Scores:

Pred. No.: 0.0806

Score: 91.50

Percent Similarity: 35.64%

Best Local Similarity: 22.77%

Query Match: 8.18%

DB: 9

Length: 3265

Matches: 46

Conservative: 26

Mismatches: 55

Indels: 75

Gaps: 8

US-10-034-500-2 (1-218) x US-10-175-752-257 (1-3265)

```

Oy 17 LeuValAspTySerGlySerMetMetLeuYHisValAlaValArgGluProLysIle 36
Db 992 GTCTTGTATAGTCTGGAAGCATGGG----- 978
Oy 37 GluLeuAlaValGluAlaIleLeuLysIleAsnAlaAlaMetProLysMetSerTyrgln 56
Db 979 -----GGTAAGACCCGCTTAATCGAATGAATCAACACAGCAAAACATTCCTGTCAG 1032
Oy 57 GlyGlyLeuTyrrThrPheAlaProTySerValIleIleProGlnGlySerTrpAsnSer 76
Db 1033 -----ACTGTTGAAAATGAAATCTCGGTGGGG 1059
Oy 77 CysVal-----AlaGlu 80
Db 1060 ATGGTTCACTTGTATGACTGCCACTATTGTAAATAGCTAATCAATAAAAGCAGT 1119
Oy 81 CysAlaValAanThrIleLeuSerAspLeuGluIlePheGlyArgLeuThrProValGly 1100
Db 1120 GATGAAAGAAACACACTGCTGACAGATTACCTACATAT-----CCTCTGGGA 1167
Oy 101 -----AspGlyIleLeuSerHisGluThrValIleAsnGlnMetProPro 115
Db 1168 GGAACCTTCATCTGCTCGTGAATTAATATGCAATTCAGGTGATTGAGAGCACTACATTC 1227
Oy 116 Gln-----AlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsn 132
Db 1228 CAACTCGATGATCCGAACTGCTGCTGACTGATGGGAGATAAACCTGCAAGTTCT 1287
Oy 133 ProValGluGluValIleYserIleTyrglnThrAsnProAsnValCysPheHisValVal 152
Db 1288 TGTATTGATGAAGTGAAA-----CNAAGTGGGGCCATTGTTTATTGCTTTG 1338
Oy 153 SerPheAla----- 155
Db 1339 GGAAGAGCTGCTGATGAACAGTAATAGATGAGCAAGATAACAGAGAGTCAATTTT 1398
Oy 156 -----AspAspAlaGluGlyLysAlaIleIleAspGlnIleValAlaLeuAsnSer 172
Db 1399 TATGTTTCAGATGAAGCTCAGAACATGGCCTCATTTGATGCTTTGGGGCTCTTACATCA 1458
Oy 173 GlySer 174
Db 1459 GGAAT 1464

```

## RESULT 25

US-10-176-482-257

; Sequence 257, Application US/10176482

; Publication No. US20030022296A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430R1C70

; CURRENT APPLICATION NUMBER: US/10/176,482

; CURRENT FILING DATE: 2002-06-20

; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 257

; LENGTH: 3265

; TYPE: DNA

; ORGANISM: Homo Sapien

US-10-176-482-257

**Alignment Scores:**

Pred. No.:	0.0806	length:	3265
Score:	91.50	Matches:	46
Percent Similarity:	33.64	Conservative:	26
Best Local Similarity:	22.77	Mismatches:	55
Query Match:	8.18	Indels:	75
DB:	9	Gaps:	8

US-10-034-500-2 (1-218) x US-10-176-482-257 (1-3265)

Qy	17	LeuValAspTyrSerGlySerMetMetMetLeuValAlaValAlaArgLysProValIle	36
Db	952	GTTCTTGATAGTCTGGAAACATGGG-----	978
Qy	37	GluLeuAlaIleValGluAlaIleLeuValIleAsnAlaAlaMetProLysMetSerTyrGln	56
Db	979	-----GGTAAAGACCCGCTAAATTCGATGAATCAAGACGAAACATTTCTGCTGCAG	10322
Qy	57	GlyGlyLeuTyrThrPheAlaProTyrSerValIleIleProGlnGlySerThrAsnSer	76
Db	1033	-----ACTGTGAAAATGATCTGGGTGGG3	1059
Qy	77	CysVal-----AlaGlu	80
Db	1060	ATGGTTCACCTTGATAGTACTGCGACATTTGTAAATAGCTAATCCAAATAAAGACAGT	1119
Qy	81	CysAlaValAsnThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGly	100
Db	1120	GATGAAGAAACACACCTCAGCGAGGATTACTACATAT-----CCTTGGGA	1167
Qy	101	-----AspGlyIleLysMetHisGluThrValIleAsnGlnMetProPro	115
Db	1168	GGAACCTTCATCTGCTCTGGAAATTTAAATATGCAATTTCAAGGTGATGGAGAGCTACATTC	1227
Qy	116	Gln-----AlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsn	132
Db	1228	CAATCGATGATCGGAAGTACGTGCTGCTGATCGATGGGGAAGATTAACATGSCAAGTTCT	1287
Qy	133	ProValGluGluValLysSerIleTyrGlnThrAsnProAsnValCysPheHisValVal	152
Db	1288	TGTTATGATGAAGTGA-----CAAAAGGGGCGCATTTGTTATTTATGCTTTG	1338
Qy	153	SerPheAla-----	155
Db	1339	GGAAGAGCTCTGATGAAAGCATTAATAGATGAGCAAGATTAACAGAGAACTATT	1398
Qy	156	-----AspAspAlaGluGlyLysValAlaIleIleAspGlnIleValAlaLeuAsnSer	172
Db	1399	TATGTTTCAGATGAAGCTCAGAACATAGGCGCTATGATGCTTTGGGCGCTTACATCA	1458
Qy	173	GlySer	174
Db	1459	GGAAAT	1464

RESULT 26  
US-10-176

US-10-176-757-257

; Sequence 257, Application US/10176757  
; Publication No. US20030022297A1  
; Attorney: TURTLEBERRY

**GENERAL INFORMATION:**

APPLICANT: Baker, Kevin P.

APPLICANT: Chen, Jian

APPLICANT: Desnoyers, Luc

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Aublin L.

APPLICANT: Smith, James  
APPLICANT: Smith, Victoria

APPLICANT: Watanabe Col in P

APPLICANT: Wood William T

APPLICANT: Zhang Zemin

TITLE OF INVENTION: SECRETET

TITLE OF INVENTION: ACIDS

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

;; FILE REFERENCE: P3430R1C86  
;; CURRENT APPLICATION NUMBER

; CURRENT FILING DATE: 2002-06-20

; Prior Application removed - See File Wrapper or Paine

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 257

; LENGTH: 3265

TYPE: DNA

! ORGANISM: Homo sapiens  
ME-10-176-757-357

CS-151-9/T-0T-50

**Alignment Scores:**

Pred. No.:	0.0860	3265
Score:	91.50	46
Percent Similarity:	35.64%	Conservative: 26
Best Local Similarity:	22.77%	Mismatches: 55
Query Match:	8.18%	Indels: 75
DB:	9	Gaps: 8

US-10-034-500-2 (1-218) x US-10-176-757-257 (1-3265)

Oy	17	LeuValAlaPyrSerGlySerMetMetCysHisValAlaValArgIleProlysile	36
Db	952	GTTCTTATTAAGCTCGAAGCATGGGG-----	978
Oy	37	GluLeuAlaLysGluAlaIleLeuLysIleAsnAlaAlaMetProLysMetSerTyrlin	56
Db	979	-----GGTAGACGCCGCTAAATCGAATGAATCAAGCAAACATTTCCGTCTGCAG	1032
Oy	57	GlyIleuLeuThrPheAlaProTyrSerValIleIleProGlnGlySerTrpAsnser	76
Db	1033	-----ACTGTTGAATAATGATCTCGGTGGG	1059
Oy	77	CysVal-----AlaGlu	80
Db	1060	ATGGTTCACTTGATAGTACTGCCACTATTGTGAATTAAGCTAATTCGAATTAAGAAGCAGT	1119
Oy	81	CysAlaValAsnThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValcyl	100
Db	1120	GATGAAGAACAACACACTGACGTGAGGATTAACCTACATAT-----CCTCTGGGA	1167
Oy	101	-----AspGlyIleLysMetHisGluThrValIleAsnGlnMetProPro	115
Db	1168	GAAACTTCATCTGCTCTGGAAATTAATATGCAATTTCAGGTGATTTGGAGACTCAATTC	1227
Oy	116	Gln-----AlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsn	132
Db	1228	CAACTGATGATGCCGAAGTACGTGCTGCTGATGATGGGAGAGATAACACTGCCAAGTCT	1287
Oy	133	ProValGluGluValLysSerIleTyrGlnThrAsnProAsnValCysPheHisValVal	152
Db	1288	TGTTTATGATGAATGAAA-----CAAGTGGGGCCATTTTCATTTATTGCTTGG	1338
Oy	153	SerPheAla-----	155
Db	1339	GGAAGAGCTCTGATAGACAGATATAGATAGACAGATTAACGAGGAAGCATTTT	1398
Oy	156	-----AspAspAlaGluGlyLysAlaIleIleAspGlnIleValAlaLeuAsnSer	172
Db	1399	TATGTTTCAGATGAAGCTCAGAAACAATGCGCTTCATTTGATGCTTTGGGCTCTTATCATCA	1458
Oy	173	GlySer	174
Db	1459	GGAAAT	1464

**RESULT 27**

US-10-176-913-257

; sequence 25/, Application US/101/6913  
; Publication No. US20030022298A1

; GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Chen, Jian

APPLICANT: Desnoyers, Luc

APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3430R1C66  
CURRENT APPLICATION NUMBER: US/10/176,913  
PRIORITY FILING DATE: 2002-06-20  
Prior Application removed - See file wrapper or Palm  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 257  
LENGTH: 3265  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-176-913-257

Alignment Scores:  
Pred. No.: 0.0806 Length: 3265  
Score: 91.50 Matches: 46  
Percent Similarity: 35.64% Conservative: 26  
Best Local Similarity: 22.77% Mismatches: 55  
Query Match: 8.18% Indels: 75  
DB: 9 Gaps: 8

US-10-034-500-2 (1-218) x US-10-176-913-257 (1-3265)

Qy 17 LeuValAspTyrSerGlySerMetMetLysHisValAlaValArgGluProLysIle 36  
Db 952 GTTCTTGAATAGCTCGAAGCATGGG----- 978  
Qy 37 GluLeuAlaValGluAlaIleLeuLysIleAsnAlaIleMetProLysMetSerTyrGln 56  
Db 979 -----GGTAAGACCGCCTTAATGAAATCAAGCAAGAAACATTTCTGCTGCAG 1032  
Qy 57 GlyGlyLeuTyrThrPheAlaProTyrSerValIleIleProGlnGlySerTyrPheAsn 76  
Db 1033 -----ACTGTTGAAATGATCTGGGTGGG 1059  
Qy 77 CysVal-----AlaGlu 80  
Db 1060 ATGGTTCACCTTGTAGTACTGCGCACTATTGTAATTAAGCTAATCAATAAAGCACT 1119  
Qy 81 CysAlaValAsnThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGly 100  
Db 1120 GATGAAGAAACACACTCATGCGAGATTACCTACATAT-----CCTCTGGGA 1167  
Qy 101 -----AspGlyIleLysMetHisGluThrValIleAsnGlnMetProPro 115  
Db 1168 GGAATTCATCTGCTGCTGAATTAATATGATTTCAAGTGTGAGAGCTACATCTCC 1227  
Qy 116 Gln-----AlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsn 132  
Db 1228 CAACTGATGATCGCAAGATGATGCTGCTGCTGATGCGGAGATTAACATGCAAGTTCT 1287  
Qy 133 ProValGluGluValIlySerIleTyrGlnThrAsnProAsnValCysPheHisValVal 152  
Db 1288 TGAATTGATGAAGTGAAGAA-----CAAGTGGGGCCATTTTCATTTTATGCTTTG 1338  
Qy 153 SerPheAla----- 155  
Db 1339 GGAAGAGCTGCTGATGAGACGTAATAGAGATGAGACAGATTAACAGAGAGCACTATTTT 1398  
Qy 156 -----AspAspAlaGluGlyLysValAlaIleIleAspGlnIleValAlaIleAsnSer 172  
Db 1399 TATGTTTCAGATGAAGCTCAAGAACATGCGCTCATTTGATGCTTTTGGGGGCTTTACATCA 1458  
Qy 173 GlySer 174  
|||:::

Db 1459 GGAAT 1464

RESULT 28  
US-10-180-552-257  
Sequence 257, Application US/10180552  
Publication No. US20030022300A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3430R1C153  
CURRENT APPLICATION NUMBER: US/10/180,552  
PRIORITY FILING DATE: 2002-06-25  
Prior Application removed - See file wrapper or Palm  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 257  
LENGTH: 3265  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-180-552-257

Alignment Scores:  
Pred. No.: 0.0806 Length: 3265  
Score: 91.50 Matches: 46  
Percent Similarity: 35.64% Conservative: 26  
Best Local Similarity: 22.77% Mismatches: 55  
Query Match: 8.18% Indels: 75  
DB: 9 Gaps: 8

US-10-034-500-2 (1-218) x US-10-180-552-257 (1-3265)

Qy 17 LeuValAspTyrSerGlySerMetMetLysHisValAlaValArgGluProLysIle 36  
Db 952 GTTCTTGAATAGCTCGAAGCATGGG----- 978  
Qy 37 GluLeuAlaValGluAlaIleLeuLysIleAsnAlaIleMetProLysMetSerTyrGln 56  
Db 979 -----GGTAAGACCGCCTTAATGAAATCAAGCAAGAAACATTTCTGCTGCAG 1032  
Qy 57 GlyGlyLeuTyrThrPheAlaProTyrSerValIleIleProGlnGlySerTyrPheAsn 76  
Db 1033 -----ACTGTTGAAATGATCTGGGTGGG 1059  
Qy 77 CysVal-----AlaGlu 80  
Db 1060 ATGGTTCACCTTGTAGTACTGCGCACTATTGTAATTAAGCTAATCAATAAAGCACT 1119  
Qy 81 CysAlaValAsnThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGly 100  
Db 1120 GATGAAGAAACACACTCATGCGAGATTACCTACATAT-----CCTCTGGGA 1167  
Qy 101 -----AspGlyIleLysMetHisGluThrValIleAsnGlnMetProPro 115  
Db 1168 GGAATTCATCTGCTGCTGAATTAATATGATTTCAAGTGTGAGAGCTACATCTCC 1227  
Qy 116 Gln-----AlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsn 132  
Db 1228 CAACTGATGATCGCAAGATGATGCTGCTGCTGATGCGGAGATTAACATGCAAGTTCT 1287  
Qy 133 ProValGluGluValIlySerIleTyrGlnThrAsnProAsnValCysPheHisValVal 152  
Db 1288 TGAATTGATGAAGTGAAGAA-----CAAGTGGGGCCATTTTCATTTTATGCTTTG 1338

Qy 153 Serphea1a----- 155  
 Db 1339 GGAAGAGCTGCTGATGACAGTAATAGATGACAGATGATGACAGAGAGTCAATTT 1398  
 Qy 156 -----AaPaPa1aGlUGlYyLa1a1e1leAaPgiN1leVal1a1e1uAaNsEr 172  
 Db 1399 TATGTTTCAGATGAGCTCAGAAACAAATGCGCTCATTTGCTTTGGGCTTTACATCA 1458  
 Qy 173 Glycer 174  
 Db 1459 GGAAT 1464

RESULT 29  
 US-10-180-557-257  
 ; Sequence 257, Application US/10180557  
 ; Publication No. US20030022301A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Chen, Uian  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Pan, James  
 ; APPLICANT: Smith, Victoria  
 ; APPLICANT: Watanabe, Colin K.  
 ; APPLICANT: Wood, William I.  
 ; APPLICANT: Zhang, Zemin  
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ; FILE REFERENCE: P3430R1C147  
 ; CURRENT APPLICATION NUMBER: US/10/180,557  
 ; CURRENT FILING DATE: 2002-06-25  
 ; Prior Application removed - See File Wrapper or Palm  
 ; NUMBER OF SEQ ID NOS: 612  
 ; SEQ ID NO 257  
 ; LENGTH: 3265  
 ; TYPE: DNA  
 ; ORGANISM: Homo Sapien  
 US-10-180-557-257

Alignment Scores:  
 Pred. No.: 0.0806 Length: 3265  
 Score: 91.50 Matches: 46  
 Percent Similarity: 35.64% Conservative: 26  
 Best Local Similarity: 22.77% Mismatches: 55  
 Query Match: 8.18% Indels: 75  
 DB: Gaps: 8

US-10-034-500-2 (1-218) x US-10-180-557-257 (1-3265)  
 Qy 17 LeuVala1PtyrSerGlySerMetMetMetCysH1eVal1a1aValArgGluProlySile 36  
 Db 952 GTTCTTGTAAGTCTGGAAGCATGGGG----- 978  
 Qy 37 GluLeuAla1yGluAla1leu1ySileAa1a1aMetProlyMeSeSerTyrgin 55  
 Db 979 -----GTAAGGACCGCTTAATGATGATCAAGCAAAACATTCTCTGCTGAG 1032  
 Qy 57 GlyGlyLeuTyrrThrPheAlaProTySerVal1le1leProGlnGlySerTyraNsEr 76  
 Db 1033 -----ACGTGTAATAATGATCCTGGGTGGGG 1059  
 Qy 77 CysVal-----AlaGlu 80  
 Db 1060 ATGGTTCACCTTGATAGTACGCCACTATTGTAATAAGCTAATCAATAAAGACAGT 1119  
 Qy 81 CysAlaVala1eNth1le1ySeSaPleuGlu1lePheGlyArgLeuThrProValGly 100  
 Db 1120 GATGAAGAACAACACTCATGCAAGATTACTACATAT-----CTCTGGGA 1167  
 Qy 101 -----AaPgi1le1yMeTh1eGluThrVal1leAaNsGlnMeCPPro 115

Db 1168 GGAATTCATCTGCTCTGGAATTAATATGATTCAGGTGATTCGAGACTACATTC 1227  
 Qy 116 Gln-----Ala1a1aVal1le1e1u1eThrAaPgiN1h1SaNsAaNsGlyMeSaN 132  
 Db 1228 CACTCGATGATTCGGAAGTACTGCTGCTGATGGGAGAGATACACTGCAAGTTCT 1287  
 Qy 133 ProVal1GluGluAla1ySeSer1leTyrg1NthraNsProAaNsVal1CysPheH1eVal1a1 152  
 Db 1288 TGTATGTATGATGAAGTGA-----CAAGTGGGGCCATTGTCATTTATTGCTTTG 1338  
 Qy 153 Serphea1a----- 155  
 Db 1339 GGAAGAGCTGCTGATGACAGTAATAGATGACAGATGATGACAGAGAGTCAATTT 1398  
 Qy 156 -----AaPaPa1aGlUGlYyLa1a1e1leAaPgiN1leVal1a1e1uAaNsEr 172  
 Db 1399 TATGTTTCAGATGAGCTCAGAAACAAATGCGCTCATTTGCTTTGGGCTTTACATCA 1458  
 Qy 173 Glycer 174  
 Db 1459 GGAAT 1464

RESULT 30  
 US-09-990-438-378  
 ; Sequence 378, Application US/09990438  
 ; Publication No. US20030027754A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ashkenazi, Avi J.  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Botstein, David  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Baton, Dan L.  
 ; APPLICANT: Ferrara, Napoleone  
 ; APPLICANT: Fong, Sherman  
 ; APPLICANT: Gerber, Hanspeter  
 ; APPLICANT: Gerritsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Grimaldi, J. Christopher  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Kijavits, Ivar J.  
 ; APPLICANT: Napier, Mary A.  
 ; APPLICANT: Pan, James  
 ; APPLICANT: Paoni, Nicholas F.  
 ; APPLICANT: Roy, Margaret Ann  
 ; APPLICANT: Stewart, Timothy A.  
 ; APPLICANT: Tumas, Daniel  
 ; APPLICANT: Watanabe, Colin K.  
 ; APPLICANT: Williams, P. Mickey  
 ; APPLICANT: Wood, William I.  
 ; APPLICANT: Zhang, Zemin  
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 ; FILE REFERENCE: P2730P1C3  
 ; CURRENT APPLICATION NUMBER: US/09/990,438  
 ; CURRENT FILING DATE: 2001-11-14  
 ; Prior Application Number: 60/049787  
 ; Prior Application Number: 60/062250  
 ; Prior Application Number: 60/062250  
 ; Prior Application Number: 60/062250  
 ; Prior Application Number: 60/065186  
 ; Prior Application Number: 60/065186  
 ; Prior Application Number: 60/065311  
 ; Prior Application Number: 60/066770  
 ; Prior Application Number: 60/066770  
 ; Prior Application Number: 60/075945  
 ; Prior Application Number: 60/075945  
 ; Prior Application Number: 60/078910  
 ; Prior Application Number: 60/083322  
 ; Prior Application Number: 60/083322  
 ; Prior Application Number: 60/084600



[illegible]

; PRIOR FILING DATE: 1998-07-07  
 ; PRIOR APPLICATION NUMBER: 60/091982  
 ; PRIOR FILING DATE: 1998-07-07  
 ; PRIOR APPLICATION NUMBER: 60/092182  
 ; PRIOR FILING DATE: 1998-07-09  
 ;

Alignment Scores:	
Pred. No.:	0.0806
Score:	91.50
Percent Similarity:	35.66%
Best Local Similarity:	22.77%
Query Match:	8.18%
DB:	9
	8

US-10-034-500-2 (1-218) X US-09-990-438-378 (1-3265)

Qy	17	LeuValaAerPyrSerGlySerMetMetLeuValaValaArgIleProLysIle	36
Db	952	GTTCCTGATTAAGCTGGAAACATGGGG	978
Qy	37	GluLeuAlaLysGluAlaIleLeuLysIleAsnAlaIleMetProLysMetSerTyGln	56
Db	979	-----GGTAAGACCGCCCTTAATTCGAATGAATCAAGCAAGCAAAACATTTCCTGCTGCAG	1032
Qy	57	GlyGlyLeuTyThrThrPheAlaProTyrSerValIleIleProGlnGlySerThrAsnSer	76
Db	1033	-----ACTGTGAAAATGGATCTGGGGTGGGG	1058
Qy	77	CysVal	80
Db	1060	ATGGTTCACCTTGTGATAGTACTGCCACATTTGTAAATTAAGCTAATCAATTAATAAGCAGT	1119
Qy	81	CysAlaValaAsnThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValaGly	100
Db	1120	GATGAAGAAGAACACACTCAGTGGAGGATTAACCTACATAT	1167
Qy	101	-----AspGlyIleLysMetHisGluThrValIleAsnGlnMetProPro	115
Db	1168	GGAACTTCACATCGCTGCTGTGGAAATTAATATGCACTTTCAGGTGATTTGGAGAGCTACATTC	1227
Qy	116	Gln-----AlaAlaValaIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsn	132
Db	1228	CAATCGATGAAGATCCGAAAGTACTGCTGCTGACTGATGGAGGAGGATTAACACTGCAAGTCT	1287
Qy	133	ProValaGluGluValLysSerIleTyArgInThrAsnProAsnValCysPheHisValaVal	152
Db	1288	TGTATTTGATGAAGTGAAGAA-----CAAAAGTGGGGCCATTGTTCAATTTATTTGCTTGG	1338
Qy	153	SerPheAla	155
Db	1339	GGAAAGACTCTGTATGAACGATTAATAGATAGTACGAAGTAAACAGGAAGTCAATTT	1398
Qy	156	-----AspAspAlaGluGlyLysValaIleIleAspGlnIleValaAlaLeuAsnSer	172
Db	1399	TATGTTTCAGATGAAGCTCAGAACAAATGGCCTCATTTGATGCTTTGGGCTCTTACATCA	1458
Qy	173	GlySer	174
Db	1459	GGAAAT	1464

```

RESULT 31
US-09-990-562-378
: Sequence 378, Application US/09990562Z
: Publication No. US20030027985A1
: GENERAL INFORMATION:
: APPLICANT: Ashkenazi, Avi J.
: APPLICANT: Baker, Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Desnovers, Luc
: APPLICANT: Eaton, Dan L.
: APPLICANT: Ferrara, Napoleone
: APPLICANT: Fong, Sherman
: APPLICANT: Gettel, Hanspeter

```

APPLICANT: Gerltsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gueney, Austin L.  
APPLICANT: Kijavini, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
TITLE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: P2730P18  
CURRENT APPLICATION NUMBER: US/09/990,562  
CURRENT FILING DATE: 2001-11-14  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05

1	PRIOR APPLICATION NUMBER: 60/088655
2	PRIOR FILING DATE: 1998-06-09
3	PRIOR APPLICATION NUMBER: 60/088734
4	PRIOR FILING DATE: 1998-06-10
5	PRIOR APPLICATION NUMBER: 60/088738
6	PRIOR FILING DATE: 1998-06-10
7	PRIOR APPLICATION NUMBER: 60/088742
8	PRIOR FILING DATE: 1998-06-10
9	PRIOR APPLICATION NUMBER: 60/088810
10	PRIOR FILING DATE: 1998-06-10
11	PRIOR APPLICATION NUMBER: 60/088824
12	PRIOR FILING DATE: 1998-06-10
13	PRIOR APPLICATION NUMBER: 60/088826
14	PRIOR FILING DATE: 1998-06-10
15	PRIOR APPLICATION NUMBER: 60/088858
16	PRIOR FILING DATE: 1998-06-11
17	PRIOR APPLICATION NUMBER: 60/088861
18	PRIOR FILING DATE: 1998-06-11
19	PRIOR APPLICATION NUMBER: 60/088876
20	PRIOR FILING DATE: 1998-06-11
21	PRIOR APPLICATION NUMBER: 60/089105
22	PRIOR FILING DATE: 1998-06-12
23	PRIOR APPLICATION NUMBER: 60/089440
24	PRIOR FILING DATE: 1998-06-16
25	PRIOR APPLICATION NUMBER: 60/089512
26	PRIOR FILING DATE: 1998-06-16
27	PRIOR APPLICATION NUMBER: 60/089514
28	PRIOR FILING DATE: 1998-06-16
29	PRIOR APPLICATION NUMBER: 60/089532
30	PRIOR FILING DATE: 1998-06-17
31	PRIOR APPLICATION NUMBER: 60/089538
32	PRIOR FILING DATE: 1998-06-17
33	PRIOR APPLICATION NUMBER: 60/089598
34	PRIOR FILING DATE: 1998-06-17
35	PRIOR APPLICATION NUMBER: 60/089599
36	PRIOR FILING DATE: 1998-06-17
37	PRIOR APPLICATION NUMBER: 60/089600
38	PRIOR FILING DATE: 1998-06-17
39	PRIOR APPLICATION NUMBER: 60/089653
40	PRIOR FILING DATE: 1998-06-17
41	PRIOR APPLICATION NUMBER: 60/089801
42	PRIOR FILING DATE: 1998-06-18
43	PRIOR APPLICATION NUMBER: 60/089907
44	PRIOR FILING DATE: 1998-06-18
45	PRIOR APPLICATION NUMBER: 60/089908
46	PRIOR FILING DATE: 1998-06-18
47	PRIOR APPLICATION NUMBER: 60/089947
48	PRIOR FILING DATE: 1998-06-19
49	PRIOR APPLICATION NUMBER: 60/089948
50	PRIOR FILING DATE: 1998-06-19
51	PRIOR APPLICATION NUMBER: 60/090252
52	PRIOR FILING DATE: 1998-06-22
53	PRIOR APPLICATION NUMBER: 60/090254
54	PRIOR FILING DATE: 1998-06-22
55	PRIOR APPLICATION NUMBER: 60/090349
56	PRIOR FILING DATE: 1998-06-23
57	PRIOR APPLICATION NUMBER: 60/090355
58	PRIOR FILING DATE: 1998-06-23
59	PRIOR APPLICATION NUMBER: 60/090429
60	PRIOR FILING DATE: 1998-06-24
61	PRIOR APPLICATION NUMBER: 60/090431
62	PRIOR FILING DATE: 1998-06-24
63	PRIOR APPLICATION NUMBER: 60/090435
64	PRIOR FILING DATE: 1998-06-24
65	PRIOR APPLICATION NUMBER: 60/090444
66	PRIOR FILING DATE: 1998-06-24
67	PRIOR APPLICATION NUMBER: 60/090455
68	PRIOR FILING DATE: 1998-06-24
69	PRIOR APPLICATION NUMBER: 60/090472

DB	Seq1	Seq2	Score	Length	Matches	Conservative	Mismatches	Indels	Gaps
US-10-034-500-2 (1-218)	x	US-09-990-562-378 (1-3265)							
Alignment Scores:									
Pred. No.:	0.0806	Length:	3265						
Score:	91.50	Matches:	46						
Percent Similarity:	35.64%	Conservative:	26						
Best Local Similarity:	22.77%	Mismatches:	55						
Query Match:	8.18%	Indels:	75						
DB:	9	Gaps:	8						
DB	17	leuValaspIlyrserGlySerMetMetMetuylshISValAlaValArgGluProLysIle	36						
DB	952	GTTCCTGATTAAGCTCGGAACATGAGG	978						
DB	37	GlueuAlalysGluIalleuLysIleAspAlaIaMetProLysMetSerTyrlin	56						
DB	979	-----GRTAAGGACCGCCCTTAATGAAATGAATCAAGCAGCAAAACATTTCTGTGTGAG	1032						
DB	57	GlyGlyLeuTyrrThrPheAlaProTyrlSerValIleIleProGlyIlySerTyrlPanser	76						
DB	1033	-----ACTGTTGAAATGATCTCGGTGGG	1059						
DB	77	CysVal-----AlaGlu	80						
DB	1060	ATGCTTCACTTGATGATGACTGCACATTGTGAAATTAAGCTAATCCAAATATAAAAGCAGT	1119						
DB	81	CysAlaValaIenThrIleLysSerSerApleuGluIlePheGlyArgLeuThrProValGly	100						
DB	1120	GATGAAGAAGAACACACTGATGCGAGGATTAACCTACATAT-----CTCTGGGA	1167						

Qy 101 -----AspGlyIyleysmethisgluthtValIleasnGlnmeCProPro 115  
Db 1168 GGAACCTTCATCTGCTCGATTAATATGATTCAGTATTCAGAGATCACTTCC 1227  
Qy 116 Gln-----AlaAlaValIleleuLeuThrAspGlyIleAsnAsnleuGlymeCAsn 132  
Db 1228 CAACCTCGATGATCCGAATGCTGCTGATGGGAGATTAACATCACTCAAGTTCT 1287  
Qy 133 ProValGluGluValIleysSerIleYrGlnThrAsnProAsnValCysPheHisValVal 152  
Db 1288 TGTATTGATGAAGTAAA-----CAAGCTGGGCGCATTTTCATTATTCCTTGG 1338  
Qy 153 SerPheAla----- 155  
Db 1339 GGAAGAGCTGCTGATGAGCACTAATAGATGACAGATTAACAGAGAGATCATTTT 1398  
Qy 156 -----AspAspAlaGluGlyIleValIleIleAspGlnIleValAlaIleuAsnSer 172  
Db 1399 TATGTTTCAGATGAAGCTCAGAACATGGCTCATTTGCTTTGGGGCTCTTACATCA 1458  
Qy 173 GlySer 174  
Db 1459 GGAAT 1464  
RESULT 32  
US-09-997-428-378  
Sequence 378, Application US/09997428  
Publication No. US20030027162A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi J.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Deenoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerliessen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Grimaldi, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zhen  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2730P1C44  
CURRENT APPLICATION NUMBER: US/09/997,428  
PRIOR FILING DATE: 2001-11-15  
PRIOR APPLICATION NUMBER: 60/045787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20

PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599

```

/ PRIOR FILING DATE: 1998-06-17
/ PRIOR APPLICATION NUMBER: 60/089600
/ PRIOR FILING DATE: 1998-06-17
/ PRIOR APPLICATION NUMBER: 60/089653
/ PRIOR FILING DATE: 1998-06-17
/ PRIOR APPLICATION NUMBER: 60/089601
/ PRIOR FILING DATE: 1998-06-18
/ PRIOR APPLICATION NUMBER: 60/089907
/ PRIOR FILING DATE: 1998-06-18
/ PRIOR APPLICATION NUMBER: 60/089908
/ PRIOR FILING DATE: 1998-06-18
/ PRIOR APPLICATION NUMBER: 60/089947
/ PRIOR FILING DATE: 1998-06-19
/ PRIOR APPLICATION NUMBER: 60/089948
/ PRIOR FILING DATE: 1998-06-19
/ PRIOR APPLICATION NUMBER: 60/089952
/ PRIOR FILING DATE: 1998-06-19
/ PRIOR APPLICATION NUMBER: 60/090246
/ PRIOR FILING DATE: 1998-06-22
/ PRIOR APPLICATION NUMBER: 60/090252
/ PRIOR FILING DATE: 1998-06-22
/ PRIOR APPLICATION NUMBER: 60/090254
/ PRIOR FILING DATE: 1998-06-22
/ PRIOR APPLICATION NUMBER: 60/090349
/ PRIOR FILING DATE: 1998-06-23
/ PRIOR APPLICATION NUMBER: 60/090355
/ PRIOR FILING DATE: 1998-06-23
/ PRIOR APPLICATION NUMBER: 60/090429
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090431
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090435
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090444
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090445
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090472
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090535
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090540
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090542
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090557
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090676
/ PRIOR FILING DATE: 1998-06-25
/ PRIOR APPLICATION NUMBER: 60/090678
/ PRIOR FILING DATE: 1998-06-25
/ PRIOR APPLICATION NUMBER: 60/090690
/ PRIOR FILING DATE: 1998-06-25
/ PRIOR APPLICATION NUMBER: 60/090694
/ PRIOR FILING DATE: 1998-06-25
/ PRIOR APPLICATION NUMBER: 60/090695
/ PRIOR FILING DATE: 1998-06-25
/ PRIOR APPLICATION NUMBER: 60/090696
/ PRIOR FILING DATE: 1998-06-25
/ PRIOR APPLICATION NUMBER: 60/090862
/ PRIOR FILING DATE: 1998-06-26
/ PRIOR APPLICATION NUMBER: 60/090863
/ PRIOR FILING DATE: 1998-06-26
/ PRIOR APPLICATION NUMBER: 60/091360
/ PRIOR FILING DATE: 1998-07-01
/ PRIOR APPLICATION NUMBER: 60/091478
/ PRIOR FILING DATE: 1998-07-02
/ PRIOR APPLICATION NUMBER: 60/091544
/ PRIOR FILING DATE: 1998-07-01
/ PRIOR APPLICATION NUMBER: 60/091519
/ PRIOR FILING DATE: 1998-07-02
/ PRIOR APPLICATION NUMBER: 60/091626
/ PRIOR FILING DATE: 1998-07-02

```

```

/ PRIOR APPLICATION NUMBER: 60/091633
/ PRIOR FILING DATE: 1998-07-02
/ PRIOR APPLICATION NUMBER: 60/091978
/ PRIOR FILING DATE: 1998-07-07
/ PRIOR APPLICATION NUMBER: 60/091982
/ PRIOR FILING DATE: 1998-07-07
/ PRIOR APPLICATION NUMBER: 60/092182
/ PRIOR FILING DATE: 1998-07-09

```

## Alignment Scores:

Pred. No.:	0	0806	Length:	3265
Score:	91.50	35.64%	Matches:	46
Best Local Similarity:	22.77%		Mismatches:	55
Query Match:	8.18%		Indels:	75
DB:	9		Gaps:	8

US-10-034-500-2 (1-218) x US-09-997-428-378 (1-3265)

```

Qy 17 LeuValAspTyrSerGlySerMetMetLeuYshValAlaValArgGluProLysIle 36
      :::::|||||
Db 952 GTTCTTGAATAGTCTGGAAGCATGGG----- 978
Qy 37 GluLeuAlaLysGluAlaIleLeuLysIleAsnAlaIleMetProLysMetSerTyrGln 56
      ||:::|||||
Db 979 -----GGTAAGACCGCCTTAATGAAATGAAATCAAGACAAAATTTCCTGCTGAG 1032
Qy 57 GlyLysLeuTyrThrPheAlaProTyrSerValIleIleProGlnGlySerTyrAsnSer 76
      :::::|||||
Db 1033 -----ACTGTTGAAATGATCTCGGTGGG 1059
Qy 77 CysVal-----AlaGlu 80
      |||
Db 1066 ATGGTTCACTTGATAGTACGTCACATATTGTAATAAGCTAATCAATAAAGCAGT 1119
Qy 81 CysAlaValAsnThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGly 100
      ||:::|||||
Db 1120 GATGAAGAAACACACTGATGACAGATTACCTACATAT-----CCTCTGGA 1167
Qy 101 -----AspGlyIleLysMetHisGluThrValIleAsnGlnMetProPro 115
      |||
Db 1168 GGAATCTTCATGCTGCTGGAATTAATATGATTCATTCAGTGATTTGAGAGCTTACATTC 1227
Qy 116 Gln-----AlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsn 132
      |||
Db 1228 CAATCGATGATCGAAGTACTGCTGCTGATGATGGGAGATTAACATCAAGTCT 1287
Qy 133 ProValGluGluValLysSerIleTyrGlnThrAsnProAsnValCysPheHisValVal 152
      :::::|||||
Db 1288 TGTATTGATGAAGTAAA-----CAAGTGGGGCCATTGTTCAATTTATTCCTTGG 1338
Qy 153 SerPheAla----- 155
      |||
Db 1339 GGAAGAGCTGCTGATGACCAATATAGATGACAGACAAATAACAGAGAGATCATTTT 1398
Qy 156 -----AspAspAlaGluGlyLysAlaIleIleAspGlnIleValAlaLeuAsnSer 172
      |||
Db 1399 TATGTTTGAATGAAGCTCAGAAACATGGCCCTCATTTGATGCTTTGGGCTCTTACATCA 1458
Qy 173 GlySer 174
      |||
Db 1459 GGAAT 1464

```

## RESULT 33

```

US-09-997-666-378
/ Sequence 378, Application US/09997666
/ Publication No. US20030027163A1

```

## GENERAL INFORMATION:

```

/ APPLICANT: Ashkenazi, Avi J.
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Botstein, David
/ APPLICANT: Desnovers, Luc
/ APPLICANT: Eaton, Dan L.

```

APPLICANT: Ferrara, Napoleone  
 APPLICANT: Fong, Sherman  
 APPLICANT: Gerber, Hanspeter  
 APPLICANT: Gerritsen, Mary E.  
 APPLICANT: Goddard, Audrey  
 APPLICANT: Godowski, Paul J.  
 APPLICANT: Grimaldi, J. Christopher  
 APPLICANT: Gunney, Austin L.  
 APPLICANT: Klavin, Ivar J.  
 APPLICANT: Napier, Mary A.  
 APPLICANT: Pan, James  
 APPLICANT: Paoni, Nicholas F.  
 APPLICANT: Roy, Margaret Ann  
 APPLICANT: Stewart, Timothy A.  
 APPLICANT: Tumas, Daniel  
 APPLICANT: Macanabe, Colin K.  
 APPLICANT: Williams, P. Mickey  
 APPLICANT: Wood, William I.  
 APPLICANT: Zhang, Zemin  
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 FILE REFERENCE: P2730P1C42  
 CURRENT APPLICATION NUMBER: US/09/997,666  
 PRIOR FILING DATE: 2001-11-15  
 PRIOR APPLICATION NUMBER: 60/043787  
 PRIOR FILING DATE: 1997-06-16  
 PRIOR APPLICATION NUMBER: 60/062250  
 PRIOR FILING DATE: 1997-10-17  
 PRIOR APPLICATION NUMBER: 60/065186  
 PRIOR FILING DATE: 1997-11-12  
 PRIOR APPLICATION NUMBER: 60/065311  
 PRIOR FILING DATE: 1997-11-13  
 PRIOR APPLICATION NUMBER: 60/066770  
 PRIOR FILING DATE: 1997-11-24  
 PRIOR APPLICATION NUMBER: 60/075945  
 PRIOR FILING DATE: 1998-02-25  
 PRIOR APPLICATION NUMBER: 60/078910  
 PRIOR FILING DATE: 1998-03-20  
 PRIOR APPLICATION NUMBER: 60/083322  
 PRIOR FILING DATE: 1998-04-28  
 PRIOR APPLICATION NUMBER: 60/084600  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/087106  
 PRIOR FILING DATE: 1998-05-28  
 PRIOR APPLICATION NUMBER: 60/087607  
 PRIOR FILING DATE: 1998-06-02  
 PRIOR APPLICATION NUMBER: 60/087609  
 PRIOR FILING DATE: 1998-06-02  
 PRIOR APPLICATION NUMBER: 60/087759  
 PRIOR FILING DATE: 1998-06-02  
 PRIOR APPLICATION NUMBER: 60/087827  
 PRIOR FILING DATE: 1998-06-03  
 PRIOR APPLICATION NUMBER: 60/088021  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088025  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088026  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088028  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088029  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088030  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088033  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088326  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088167  
 PRIOR FILING DATE: 1998-06-05  
 PRIOR APPLICATION NUMBER: 60/088202  
 PRIOR FILING DATE: 1998-06-05  
 PRIOR APPLICATION NUMBER: 60/088212

PRIOR FILING DATE: 1998-06-05  
 PRIOR APPLICATION NUMBER: 60/088217  
 PRIOR FILING DATE: 1998-06-05  
 PRIOR APPLICATION NUMBER: 60/088655  
 PRIOR FILING DATE: 1998-06-09  
 PRIOR APPLICATION NUMBER: 60/088734  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088738  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088742  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088810  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088824  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088826  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088858  
 PRIOR FILING DATE: 1998-06-11  
 PRIOR APPLICATION NUMBER: 60/088861  
 PRIOR FILING DATE: 1998-06-11  
 PRIOR APPLICATION NUMBER: 60/088876  
 PRIOR FILING DATE: 1998-06-11  
 PRIOR APPLICATION NUMBER: 60/089105  
 PRIOR FILING DATE: 1998-06-12  
 PRIOR APPLICATION NUMBER: 60/089440  
 PRIOR FILING DATE: 1998-06-16  
 PRIOR APPLICATION NUMBER: 60/089512  
 PRIOR FILING DATE: 1998-06-16  
 PRIOR APPLICATION NUMBER: 60/089514  
 PRIOR FILING DATE: 1998-06-16  
 PRIOR APPLICATION NUMBER: 60/089532  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089538  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089598  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089599  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089600  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089653  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089801  
 PRIOR FILING DATE: 1998-06-18  
 PRIOR APPLICATION NUMBER: 60/089907  
 PRIOR FILING DATE: 1998-06-18  
 PRIOR APPLICATION NUMBER: 60/089908  
 PRIOR FILING DATE: 1998-06-18  
 PRIOR APPLICATION NUMBER: 60/089947  
 PRIOR FILING DATE: 1998-06-19  
 PRIOR APPLICATION NUMBER: 60/089948  
 PRIOR FILING DATE: 1998-06-19  
 PRIOR APPLICATION NUMBER: 60/089952  
 PRIOR FILING DATE: 1998-06-19  
 PRIOR APPLICATION NUMBER: 60/090246  
 PRIOR FILING DATE: 1998-06-22  
 PRIOR APPLICATION NUMBER: 60/090252  
 PRIOR FILING DATE: 1998-06-22  
 PRIOR APPLICATION NUMBER: 60/090254  
 PRIOR FILING DATE: 1998-06-22  
 PRIOR APPLICATION NUMBER: 60/090349  
 PRIOR FILING DATE: 1998-06-23  
 PRIOR APPLICATION NUMBER: 60/090355  
 PRIOR FILING DATE: 1998-06-23  
 PRIOR APPLICATION NUMBER: 60/090429  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090431  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090435  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090444  
 PRIOR FILING DATE: 1998-06-24

```

PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090662
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

```

## Alignment Scores:

```

Pred. No.: 0.0806 Length: 3265
Score: 91.50 Matches: 46
Percent Similarity: 35.64% Conservative: 26
Best Local Similarity: 22.77% Mismatches: 55
Query Match: 8.18% Indels: 75
DB: 9 Gaps: 8

```

US-10-034-500-2 (1-218) x US-09-997-666-378 (1-3265)

```

QY 17 LeuValAspTyrSerGlySerMetMetMetLysHisValAlaValArgGluProLysIle 36
DB 952 GTTCTGATGAAGCTGGAACATGGGG----- 978
QY 37 GluLeuAlaIysGluAlaIleLeuLysIleAsnAlaIleMetProLysMetSerTyrGln 56
DB 979 -----GGTAAGAGCCGCTTAATGCATGATCAACAGCAAAACATTTCCGCTGAG 1032
QY 57 GlyGlyLeuTyrThrPheAlaProTyrSerValIleIleProGlnIysSerTyrPasnSer 76
DB 1033 -----ACTGTGAATAATGATCTCGGCGG 1059
QY 77 CysVal-----AlaGlu 80
DB 1060 ATGTTCACTTGTATAGTACTGCCACTATTGTAAATAGCTAATCAAAATAAAGCAGT 1119

```

```

QY 81 CysAlaValAsnThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGly 100
DB 1120 GATGAAGAAACACACTCATCGCAGGATTAACCTACATAT-----CCTCTGGGA 1167
QY 101 -----AspGlyIleLysMetHisGluThrValIleAsnGlnMetProPro 115
DB 1168 GGAACCTTCATCTGCTGTGAATTAAATATGCATTTCAGGGATGGAGACTTCATTCC 1227
QY 116 Gln-----AlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsn 132
DB 1228 CAATCGATGATCGGAGTACTGCTGCTGACTGATGGGAGGATTAACACGCAAGTTCT 1287
QY 133 ProValGluGluValLysSerIleTyrGlnThrAsnProAsnValCysPheHisValVal 152
DB 1288 TGTATTGATGAAGTGAAA-----CAAGTGGGCGCATTTGATTATTGCTTTG 1338
QY 153 SerPheAla----- 155
DB 1339 GGAAGAGCTGCTGATGAAGCAGTATAGATGAGCAAGATPAACAGAGAGATCATTTT 1398
QY 156 -----AspAspAlaGluGlyLysAlaIleIleAspGlnIleValAlaLeuAsnSer 172
DB 1399 TATGTTTCAAGATGAAGCTCAGAACATGCGCTCATTTGCTTTGGGCGCTTACATCA 1458
QY 173 GlySer 174
DB 1459 GGAAT 1464

```

## RESULT 34

```

US-10-173-700-257
Sequence 257, Application US/10173700
Publication No. US20030027262A1

```

## GENERAL INFORMATION:

```

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C14
CURRENT APPLICATION NUMBER: US/10/173,700
PRIOR FILING DATE: 2002-06-17
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 257
LENGTH: 3265
TYPE: DNA
ORGANISM: Homo Sapien
US-10-173-700-257

```

## Alignment Scores:

```

Pred. No.: 0.0806 Length: 3265
Score: 91.50 Matches: 46
Percent Similarity: 35.64% Conservative: 26
Best Local Similarity: 22.77% Mismatches: 55
Query Match: 8.18% Indels: 75
DB: 9 Gaps: 8

```

US-10-034-500-2 (1-218) x US-10-173-700-257 (1-3265)

```

QY 17 LeuValAspTyrSerGlySerMetMetMetLysHisValAlaValArgGluProLysIle 36
DB 952 GTTCTGATGAAGCTGGAACATGGGG----- 978
QY 37 GluLeuAlaIysGluAlaIleLeuLysIleAsnAlaIleMetProLysMetSerTyrGln 56
DB 1060 ATGTTCACTTGTATAGTACTGCCACTATTGTAAATAGCTAATCAAAATAAAGCAGT 1119

```

Db 979 -----GTTAAGACCGCTTAATGATGAATCAAGCAACAAATTTCTGCTGCAG 1032  
Qy 57 GlycylLeuTyrrhrPhealaProTyrSerValIleileProGlnGlySerTyrPanser 76  
Db 1033 -----ACTGTTGAAAATGATCTCGGGTGGGG 1059  
Qy 77 CysValI-----AlaGlu 80  
Db 1060 ATGGTTCACCTTGTGATAGTACGCCACTATTGTAAATAGCTAATCCAAATATAAAAGCAGT 1119  
Qy 81 CysAlaValaenThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGly 100  
Db 1120 GATGAAAGAAACACACTATGCGAGATTAACCTACATAT-----CCTCTGGGA 1167  
Qy 101 -----AspGlyIleLysMetHisGluThrValIleasnGlnMetProPro 115  
Db 1168 GGAACCTTCATCTGCTCGGAATTAATATGATTCAGGTATTCAGAGACTACATTC 1227  
Qy 116 Gln-----AlaAlaValIleLeuLeuThrAspGlyHisasnAsnLeuGlyMetAsn 132  
Db 1228 CAACCTCGATGATCCGAAGTACTGCTGCTGACTGATGGGAGAGATACACATGCAATTC 1287  
Qy 133 ProValGluGluValLysSerIleTyrGlnThrAsnProAsnValCysPheHisValVal 152  
Db 1288 TGTATTGATGAAGTAAA-----CAAGTGGGGCCATTGTTCAATTATTCCTTGG 1338  
Qy 153 SerPheala----- 155  
Db 1339 GGAAGAGCTGCTGATGAAGCAGTATAGAGATGACAGATACAGAGAAATCATTTT 1398  
Qy 156 -----AspAspAlaGluGlyLysAlaIleIleAspGlnIleValAlaLeuAsnSer 172  
Db 1399 TATGTTTCAGATGAAGCTCAGAACATGGCCATTCATTCGTTGGGGCTCTTACATCA 1458  
Qy 173 Glyser 174  
Db 1459 GGAAT 1464

## RESULT 35

US-10-174-572-257  
; Sequence 257, Application US/10174572  
; Publication No. US20030027263A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Deenoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3430R1C40  
; CURRENT APPLICATION NUMBER: US/10/174,572  
; CURRENT FILING DATE: 2002-06-18  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 257  
; LENGTH: 3265  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-174-572-257

## Alignment Scores:

Pred. No.: 0.0806 Length: 3265  
Score: 91.50 Matches: 46  
Percent Similarity: 35.64% Conservative: 26  
Best Local Similarity: 22.77% Mismatches: 75  
Query Match: 8.18% Indels: 75

DB: 9 Gaps: 8  
US-10-034-500-2 (1-218) x US-10-174-572-257 (1-3265)  
Qy 17 LeuValAspTyrSerGlySerMetMetLysHisValAlaValArgGluProLysIle 36  
Db 952 GTTCTGATTAAGTCTGGAAGCATGGG----- 978  
Qy 37 GluLeuAlaLysGluAlaIleLeuLysIleAsnAlaIleMetProLysMetSerTyrGln 56  
Db 979 -----GTTAAGACCGCTTAATGATGAATCAAGCAGCAACAAATTTCTGCTGCAG 1032  
Qy 57 GlycylLeuTyrrhrPhealaProTyrSerValIleileProGlnGlySerTyrPanser 76  
Db 1033 -----ACTGTTGAAAATGATCTCGGGTGGGG 1059  
Qy 77 CysValI-----AlaGlu 80  
Db 1060 ATGGTTCACCTTGTGATAGTACGCCACTATTGTAAATAGCTAATCCAAATATAAAAGCAGT 1119  
Qy 81 CysAlaValaenThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGly 100  
Db 1120 GATGAAAGAAACACACTATGCGAGATTAACCTACATAT-----CCTCTGGGA 1167  
Qy 101 -----AspGlyIleLysMetHisGluThrValIleasnGlnMetProPro 115  
Db 1168 GGAACCTTCATCTGCTCGGAATTAATATGATTCAGGTATTCAGAGACTACATTC 1227  
Qy 116 Gln-----AlaAlaValIleLeuLeuThrAspGlyHisasnAsnLeuGlyMetAsn 132  
Db 1228 CAACCTCGATGATCCGAAGTACTGCTGCTGACTGATGGGAGAGATACACATGCAATTC 1287  
Qy 133 ProValGluGluValLysSerIleTyrGlnThrAsnProAsnValCysPheHisValVal 152  
Db 1288 TGTATTGATGAAGTAAA-----CAAGTGGGGCCATTGTTCAATTATTCCTTGG 1338  
Qy 153 SerPheala----- 155  
Db 1339 GGAAGAGCTGCTGATGAAGCAGTATAGAGATGACAGATACAGAGAAATCATTTT 1398  
Qy 156 -----AspAspAlaGluGlyLysAlaIleIleAspGlnIleValAlaLeuAsnSer 172  
Db 1399 TATGTTTCAGATGAAGCTCAGAACATGGCCATTCATTCGTTGGGGCTCTTACATCA 1458  
Qy 173 Glyser 174  
Db 1459 GGAAT 1464

## RESULT 36

US-10-174-579-257  
; Sequence 257, Application US/10174579  
; Publication No. US20030027264A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Deenoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3430R1C31  
; CURRENT APPLICATION NUMBER: US/10/174,579  
; CURRENT FILING DATE: 2002-06-18  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 257  
; LENGTH: 3265



TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-174-579-257

Alignment Scores:  
Pred. No.: 0.0806 Length: 3265  
Score: 91.50 Matches: 46  
Percent Similarity: 35.64% Conservative: 26  
Best Local Similarity: 22.77% Mismatches: 55  
Query Match: 8.18% Indels: 75  
DB: 9 Gaps: 8

US-10-034-500-2 (1-218) x US-10-174-579-257 (1-3265)

QY 17 LeuValaAepTySerGlySerMetMetMetLysHisValAlaValArgLupProLysIle 36  
DB 952 GTTCTTGATAGCTCGAAGCATGGGG----- 978

QY 37 GluLeuAlaLysGluAlaIleLeuLysIleAsnAlaAlaMetProLysMetSerTyGln 56  
DB 979 -----GGTAAGACCGCCTTAATCAATCAATCAAGCAAGCAAAACATTCTCTGCTGCAG 1032

QY 57 GlyGlyLeuTyThrPheAlaProTySerValIleIleProGlnGlySerTyPheSer 76  
DB 1033 -----ACTGTTGAAATGATCTCGTGGGG 1059

QY 77 CysVal-----AlaGlu 80  
DB 1060 ATGGTTCACTTGATAGTACTGCCACTATTGTAAATACCTAATCCAAATAAAGCAGT 1119

QY 81 CysAlaValAsnThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGly 100  
DB 1120 GATGAAAGAAACACACTCATGCGAGATTAACCTACATAT-----CCTCTGGGA 1167

QY 101 -----AspGlyIleLysMetHisGluThrValIleAsnGlnMetProPro 115  
DB 1168 GGAACCTTCATCTGCTCGAATTAATATGATTCAGGTGATGGAAGCTACATTCC 1227

QY 116 Gln-----AlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsn 132  
DB 1228 CAACTCGATGATCCGAGATGCTGCTGACTGATGGGAGATTAACCTGCAAGTTCT 1287

QY 133 ProValGluGluValLysSerIleTyGlnThrAsnProAsnValCysPheHisValVal 152  
DB 1288 TGTATTGATGAAGTGA-----CAAGTGGGGCATTGTTTATTATGCTTTG 1338

QY 153 SerPheAla----- 155  
DB 1339 GGAAGAGCTGCTGATGAAGCATATAGATGAGCAAGATAACAGAGGAAGTCAATTT 1398

QY 156 -----AspAspAlaGluGlyLysAlaIleIleAspGlnIleValAlaLeuAsnSer 172  
DB 1399 TATGTTTCAGATGAGCTCGAACAATGGCCTTATGATGCTTTGGGGCTTTACATCA 1458

QY 173 GlySer 174  
DB 1459 GGAAT 1464

RESULT 37  
US-10-174-582-257  
Sequence 257, Application US/10174582  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
TITLE OF INVENTION: ACIDS ENCODING THE SAME  
FILE REFERENCE: P3430R1C36  
CURRENT APPLICATION NUMBER: US/10/174,582  
PRIORITY FILING DATE: 2002-06-18  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 257  
LENGTH: 3265  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-174-582-257

Alignment Scores:  
Pred. No.: 0.0806 Length: 3265  
Score: 91.50 Matches: 46  
Percent Similarity: 35.64% Conservative: 26  
Best Local Similarity: 22.77% Mismatches: 55  
Query Match: 8.18% Indels: 75  
DB: 9 Gaps: 8

US-10-034-500-2 (1-218) x US-10-174-582-257 (1-3265)

QY 17 LeuValaAepTySerGlySerMetMetMetLysHisValAlaValArgLupProLysIle 36  
DB 952 GTTCTTGATAGCTCGAAGCATGGGG----- 978

QY 37 GluLeuAlaLysGluAlaIleLeuLysIleAsnAlaAlaMetProLysMetSerTyGln 56  
DB 979 -----GGTAAGACCGCCTTAATCAATCAATCAAGCAAGCAAAACATTCTCTGCTGCAG 1032

QY 57 GlyGlyLeuTyThrPheAlaProTySerValIleIleProGlnGlySerTyPheSer 76  
DB 1033 -----ACTGTTGAAATGATCTCGTGGGG 1059

QY 77 CysVal-----AlaGlu 80  
DB 1060 ATGGTTCACTTGATAGTACTGCCACTATTGTAAATACCTAATCCAAATAAAGCAGT 1119

QY 81 CysAlaValAsnThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGly 100  
DB 1120 GATGAAAGAAACACACTCATGCGAGATTAACCTACATAT-----CCTCTGGGA 1167

QY 101 -----AspGlyIleLysMetHisGluThrValIleAsnGlnMetProPro 115  
DB 1168 GGAACCTTCATCTGCTCGAATTAATATGATTCAGGTGATGGAAGCTACATTCC 1227

QY 116 Gln-----AlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsn 132  
DB 1228 CAACTCGATGATCCGAGATGCTGCTGACTGATGGGAGATTAACCTGCAAGTTCT 1287

QY 133 ProValGluGluValLysSerIleTyGlnThrAsnProAsnValCysPheHisValVal 152  
DB 1288 TGTATTGATGAAGTGA-----CAAGTGGGGCATTGTTTATTATGCTTTG 1338

QY 153 SerPheAla----- 155  
DB 1339 GGAAGAGCTGCTGATGAAGCATATAGATGAGCAAGATAACAGAGGAAGTCAATTT 1398

QY 156 -----AspAspAlaGluGlyLysAlaIleIleAspGlnIleValAlaLeuAsnSer 172  
DB 1399 TATGTTTCAGATGAGCTCGAACAATGGCCTTATGATGCTTTGGGGCTTTACATCA 1458

QY 173 GlySer 174  
DB 1459 GGAAT 1464

RESULT 38  
US-10-174-588-257  
Sequence 257, Application US/10174588  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.

```

; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jlan
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C28
; CURRENT APPLICATION NUMBER: US/10/174,588
; CURRENT FILING DATE: 2002-06-18
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 257
; LENGTH: 3265
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-174-588-257

Alignment Scores:
Pred. No.: 0.0806 Length: 3265
Score: 91.50 Matches: 46
Percent Similarity: 35.64% Conservative: 26
Best Local Similarity: 22.77% Mismatches: 55
Query Match: 8.18% Indels: 75
DB: Gaps: 8

US-10-034-500-2 (1-218) x US-10-174-588-257 (1-3265)
Qy 17 LeuValAapTyRSerGlySerMetMetMetLysHisValAlaValArgGluProLysIle 36
Db 952 GTTCTTGAAGTCTGGAAGCATGGCG----- 978
Qy 37 GluLeuAlaLysGluAlaIleLeuLysIleAsnAlaAlaMetProLysMetSerTyrGln 56
Db 979 -----GGTAAGACCGCCTTAATGAATGAATCAAGCAGAAACATTTCCTGCTGCAG 1032
Qy 57 GlyGlyLeuTyThrPheAlaProTyrSerValIleIleProGlnGlySerTyrPheAsn 76
Db 1033 -----ACTGTGAATAATGATCTGGGTGGCG 1059
Qy 77 CysVal-----AlaGlu 80
Db 1060 ATGTTCACTTGTATAGTACGACCTATTGTAAATAAGCTAATCCAAATAAAAGCAGT 1119
Qy 81 CysAlaValAsnThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGly 100
Db 1120 GATGAAGAAACACACTGTCGAGATTACCTACATAT-----CCTCTGGGA 1167
Qy 101 -----AspGlyIleLysMetHisGluThrValIleAsnGlnMetProPro 115
Db 1168 GGAACCTTCATCTGCTGGAATTAATATGCAATTCAGGTGATTCGAGAGCTAATTC 1227
Qy 116 Gln-----AlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsn 132
Db 1228 CAACTCGATGATTCGCAAGTCTGCTGACTGTATGGGAGATTAACCTCAAGTTCT 1287
Qy 133 ProValGluGluValLysSerIleTyrGlnThrAsnProAsnValCysPheHisValVal 152
Db 1288 TGTATATGATGAAGTAAA-----CAAAGTGGGCGCATTTGTTTATTTGCTTTG 1338
Qy 153 SerPheAla----- 155
Db 1339 GGAAGAGCTGCTGATGAAGCAATAATAGATGAGCAAGATAACAGAGAGTCAATTT 1398
Qy 156 -----AspAspAlaGluGlyLysAlaIleIleAspGlnIleValAlaLeuAsnSer 172
Db 1399 TATGTTTCAGATGAAGCTCAGAAACAATGGCTTCATTGATGCTTTTGGGCTCTTACATCA 1458
```

```

Qy 173 GlySer 174
Db 1459 GGAAT 1464

RESULT 39
US-10-175-739-257
; Sequence 257, Application US/10175739
; Publication No. US20030027267A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jlan
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C46
; CURRENT APPLICATION NUMBER: US/10/175,739
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 257
; LENGTH: 3265
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-175-739-257

Alignment Scores:
Pred. No.: 0.0806 Length: 3265
Score: 91.50 Matches: 46
Percent Similarity: 35.64% Conservative: 26
Best Local Similarity: 22.77% Mismatches: 55
Query Match: 8.18% Indels: 75
DB: Gaps: 8

US-10-034-500-2 (1-218) x US-10-175-739-257 (1-3265)
Qy 17 LeuValAapTyRSerGlySerMetMetMetLysHisValAlaValArgGluProLysIle 36
Db 952 GTTCTTGAAGTCTGGAAGCATGGCG----- 978
Qy 37 GluLeuAlaLysGluAlaIleLeuLysIleAsnAlaAlaMetProLysMetSerTyrGln 56
Db 979 -----GGTAAGACCGCCTTAATGAATGAATCAAGCAGAAACATTTCCTGCTGCAG 1032
Qy 57 GlyGlyLeuTyThrPheAlaProTyrSerValIleIleProGlnGlySerTyrPheAsn 76
Db 1033 -----ACTGTGAATAATGATCTGGGTGGCG 1059
Qy 77 CysVal-----AlaGlu 80
Db 1060 ATGTTCACTTGTATAGTACGACCTATTGTAAATAAGCTAATCCAAATAAAAGCAGT 1119
Qy 81 CysAlaValAsnThrIleLysSerAspLeuGluIlePheGlyArgLeuThrProValGly 100
Db 1120 GATGAAGAAACACACTGTCGAGATTACCTACATAT-----CCTCTGGGA 1167
Qy 101 -----AspGlyIleLysMetHisGluThrValIleAsnGlnMetProPro 115
Db 1168 GGAACCTTCATCTGCTGGAATTAATATGCAATTCAGGTGATTCGAGAGCTAATTC 1227
Qy 116 Gln-----AlaAlaValIleLeuLeuThrAspGlyHisAsnAsnLeuGlyMetAsn 132
Db 1228 CAACTCGATGATTCGCAAGTCTGCTGACTGTATGGGAGATTAACCTCAAGTTCT 1287
Qy 133 ProValGluGluValLysSerIleTyrGlnThrAsnProAsnValCysPheHisValVal 152
```

